(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 August 2003 (07.08.2003)

PCT

(10) International Publication Number
WO 03/063688 A2

(51) International Patent Classification[7]: A61B

(21) International Application Number: PCT/US03/02500

(22) International Filing Date: 23 January 2003 (23.01.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/351,928 25 January 2002 (25.01.2002) US
60/359,903 25 February 2002 (25.02.2002) US
60/366,837 21 March 2002 (21.03.2002) US

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **LI, Joana, X.** [US/US]; 1264 Geneva Avenue, San Francisco, CA 94112 (US). **GORVAD, Ann, E.** [US/US]; 369 Marie Common, Livermore, CA 94550 (US). **CHAWLA, Narinder, K.** [US/US]; 33 Union Square, #712, Union City, CA 94587 (US). **SPRAGUE, William, W.** [US/US]; 611 13th Street # C, Sacramento, CA 95814 (US). **LEE, Soo, Yeun** [KR/US]; 40 Westdale Avenue, Daly City, CA 94015 (US). **CHANG, Hsin-Ru** [US/US]; 326 Treasure Island Drive, Belmont, CA 94002 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **KHARE, Reena** [US/US]; 12650 Orella Court, Saratoga, CA 95070 (US). **EMERLING, Brooke, M.** [US/US]; 833 W. Buena Avenue, #2108, Chicago, Il 60613 (US). **KABLE, Amy, E.** [US/US]; 2345 Polk Street #4, San Francisco, CA 94109 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **GIETZEN, Kimberly, J.** [US/US]; 691 Los Huecos Drive, San Jose, CA 95123 (US). **LEE, Sally** [US/US]; 3643 26th Street, San Francisco, CA 94110 (US). **SWARNAKAR, Anita** [CA/US]; 8 Locksley Avenue # 5D, San Francisco, CA 94122 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **WILSON, Amy, D.** [US/US]; 1056 Continentals Way #27, Belmont, CA 94002 (US). **JIN, Pei** [US/US]; 320 Curtner Avenue #D, Palo Alto, CA 94306 (US). **CHIEN, David** [US/US]; 444 Yerba Buéna Avenue, Los Altos, CA 94022 (US). **HAWKINS, Phillip, R.** [US/US]; 750 North Shorline Boulevard #115, Mountain View, CA 94043 (US). **JIANG, Xin** [US/US]; 14371 Elva Avenue, Saratoga, CA 95070 (US). **JACKSON, Alan, A.** [US/US]; 1541 Elwood Drive, Los Gatos, CA 95032 (US). **BHATIA, Umesh** [US/US]; 5212 Union Avenue, San Jose, CA 95124 (US). **BURRILL, John, D.** [US/US]; 2218 Brewster Avenue, Redwood City, CA 94062 (US). **BLAKE, Julie, J.** [US/US]; 3818 Pacheco Street, San Francisco, CA 94116 (US). **HO, Anne** [FR/US]; 1279 Poplar Avenue, #114, Sunnyvale, CA 94086 (US). **ZHENG, Wenjin** [CN/US]; 9 Sutter Creek Lane, Mountain View, CA 94043 (US). **ISON, Craig, H.** [US/US]; 1242 Weathersfield Way, San Jose, CA 95118 (US). **MARQUIS, Joseph, P.** [US/US]; 4428 Lazy Lane, San Jose, CA 95135 (US). **TRAN, Uyen, K.** [US/US]; 2638 Mabury Square, San Jose, CA 95133 (US). **LAL, Preeti, G.** [US/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **WARREN, Bridget, A.** [US/US]; 1810 S. El Camino Real #B103, Encinitas, CA 94024 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **HONCHELL, Cynthia, D.** [US/US]; 158 Laurel Street, San Carlos, CA 94070 (US). **BECHA, Shanya, D.** [US/US]; 21062 Gary Drive # 117, Castro Valley, CA 94546 (US). **LEHR-MASON, Patricia, M.** [US/US]; 360 Clarke Lane, Morgan Hill, CA 95014 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES

(57) Abstract: Various embodiments of the invention provide human protein modification and maintenance molecules (PMMM) and polynucleotides which identify and encode PMMM. Embodiments of the invention also provide expression vectors, host cells, antibodies, agonists, and antagonists. Other embodiments provide methods for diagnosing, treating , or preventing disorders associated with aberrant expression of PMMM.

WO 03/063688 A2

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

PROTEIN MODIFICATION AND MAINTENANCE MOLECULES

TECHNICAL FIELD

The invention relates to novel nucleic acids, protein modification and maintenance molecules encoded by these nucleic acids, and to the use of these nucleic acids and proteins in the diagnosis, treatment, and prevention of gastrointestinal, cardiovascular, autoimmune/inflammatory, cell proliferative, developmental, epithelial, neurological, and reproductive disorders. The invention also relates to the assessment of the effects of exogenous compounds on the expression of nucleic acids and protein modification and maintenance molecules.

BACKGROUND OF THE INVENTION

The cellular processes regulating modification and maintenance of protein molecules coordinate their function, conformation, stabilization, and degradation. Each of these processes is mediated by key enzymes or proteins such as kinases, phosphatases, proteases, protease inhibitors, isomerases, transferases, and molecular chaperones.

Kinases

Kinases catalyze the transfer of high-energy phosphate groups from adenosine triphosphate (ATP) to target proteins on the hydroxyamino acid residues serine, threonine, or tyrosine. Addition of a phosphate group alters the local charge on the acceptor molecule, causing internal conformational changes and potentially influencing intermolecular contacts. Reversible protein phosphorylation is the ubiquitous strategy used to control many of the intracellular events in eukaryotic cells. It is estimated that more than ten percent of proteins active in a typical mammalian cell are phosphorylated. Extracellular signals including hormones, neurotransmitters, and growth and differentiation factors can activate kinases, which can occur as cell surface receptors or as the activators of the final effector protein, as well as elsewhere along the signal transduction pathway. Kinases are involved in all aspects of a cell's function, from basic metabolic processes, such as glycolysis, to cell-cycle regulation, differentiation, and communication with the extracellular environment through signal transduction cascades. Inappropriate phosphorylation of proteins in cells has been linked to changes in cell cycle progression and cell differentiation. Changes in the cell cycle have been linked to induction of apoptosis or cancer. Changes in cell differentiation have been linked to diseases and disorders of the reproductive system, immune system, and skeletal muscle.

There are two classes of protein kinases. One class, protein tyrosine kinases (PTKs), phosphorylates tyrosine residues, and the other class, protein serine/threonine kinases (STKs), phosphorylates serine and threonine residues. Some PTKs and STKs possess structural characteristics of both families and have dual specificity for both tyrosine and serine/threonine

residues. Almost all kinases contain a conserved 250-300 amino acid catalytic domain containing specific residues and sequence motifs characteristic of the kinase family. (Reviewed in Hardie, G. and S. Hanks (1995) The Protein Kinase Facts Book, Vol I, Academic Press, San Diego, CA, pp. 17-20).

Phosphatases

Phosphatases hydrolytically remove phosphate groups from proteins. Phosphatases are essential in determining the extent of phosphorylation in the cell and, together with kinases, regulate key cellular processes such as metabolic enzyme activity, proliferation, cell growth and differentiation, cell adhesion, and cell cycle progression. Protein phosphatases are characterized as either serine/threonine- or tyrosine-specific based on their preferred phospho-amino acid substrate. Some phosphatases (DSPs, for dual specificity phosphatases) can act on phosphorylated tyrosine, serine, or threonine residues. The protein serine/threonine phosphatases (PSPs) are important regulators of many cAMP-mediated hormone responses in cells. Protein tyrosine phosphatases (PTPs) play a significant role in cell cycle and cell signaling processes.

Proteases

Proteases cleave proteins and peptides at the peptide bond that forms the backbone of the protein or peptide chain. Proteolysis is one of the most important and frequent enzymatic reactions that occurs both within and outside of cells. Proteolysis is responsible for the activation and maturation of nascent polypeptides, the degradation of misfolded and damaged proteins, and the controlled turnover of peptides within the cell. Proteases participate in digestion, endocrine function, tissue remodeling during embryonic development, wound healing, and normal growth. Proteases can play a role in regulatory processes by affecting the half life of regulatory proteins. Proteases are involved in the etiology or progression of disease states such as inflammation, angiogenesis, tumor dispersion and metastasis, cardiovascular disease, neurological disease, and bacterial, parasitic, and viral infections.

Proteases can be categorized on the basis of where they cleave their substrates. Exopeptidases, which include aminopeptidases, dipeptidyl peptidases, tripeptidases, carboxypeptidases, peptidyl-di-peptidases, dipeptidases, and omega peptidases, cleave residues at the termini of their substrates. Endopeptidases, including serine proteases, cysteine proteases, and metalloproteases, cleave at residues within the peptide. Four principal categories of mammalian proteases have been identified based on active site structure, mechanism of action, and overall three-dimensional structure. (See Beynon, R.J. and J.S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford University Press, New York NY, pp. 1-5.)

Serine Proteases

The serine proteases (SPs) are a large, widespread family of proteolytic enzymes that include

the digestive enzymes trypsin and chymotrypsin, components of the complement and blood-clotting cascades, and enzymes that control the degradation and turnover of macromolecules within the cell and in the extracellular matrix. Most of the more than 20 subfamilies can be grouped into six clans, each with a common ancestor. These six clans are hypothesized to have descended from at least four evolutionarily distinct ancestors. SPs are named for the presence of a serine residue found in the active catalytic site of most families. The active site is defined by the catalytic triad, a set of conserved asparagine, histidine, and serine residues critical for catalysis. These residues form a charge relay network that facilitates substrate binding. Other residues outside the active site form an oxyanion hole that stabilizes the tetrahedral transition intermediate formed during catalysis. SPs have a wide range of substrates and can be subdivided into subfamilies on the basis of their substrate specificity. The main subfamilies are named for the residue(s) after which they cleave: trypases (after arginine or lysine), aspases (after aspartate), chymases (after phenylalanine or leucine), metases (methionine), and serases (after serine) (Rawlings, N.D. and A.J. Barrett (1994) Methods Enzymol. 244:19-61).

Most mammalian serine proteases are synthesized as zymogens, inactive precursors that are activated by proteolysis. For example, trypsinogen is converted to its active form, trypsin, by enteropeptidase. Enteropeptidase is an intestinal protease that removes an N-terminal fragment from trypsinogen. The remaining active fragment is trypsin, which in turn activates the precursors of the other pancreatic enzymes. Likewise, proteolysis of prothrombin, the precursor of thrombin, generates three separate polypeptide fragments. The N-terminal fragment is released while the other two fragments, which comprise active thrombin, remain associated through disulfide bonds.

The two largest SP subfamilies are the chymotrypsin (S1) and subtilisin (S8) families. Some members of the chymotrypsin family contain two structural domains unique to this family. Kringle domains are triple-looped, disulfide cross-linked domains found in varying copy number. Kringle domains are thought to play a role in binding mediators such as membranes, other proteins or phospholipids, and in the regulation of proteolytic activity (PROSITE PDOC00020). Apple domains are 90 amino-acid repeated domains, each containing six conserved cysteines. Three disulfide bonds link the first and sixth, second and fifth, and third and fourth cysteines (PROSITE PDOC00376). Apple domains are involved in protein-protein interactions. S1 family members include trypsin, chymotrypsin, coagulation factors IX-XII, complement factors B, C, and D, granzymes, kallikrein, and tissue- and urokinase-plasminogen activators. The subtilisin family has members found in the eubacteria, archaebacteria, eukaryotes, and viruses. Subtilisins include the proprotein-processing endopeptidases kexin and furin and the pituitary prohormone convertases PC1, PC2, PC3, PC6, and PACE4 (Rawlings and Barrett, *supra*).

SPs have functions in many normal processes and some have been implicated in the etiology

or treatment of disease. Enterokinase, the initiator of intestinal digestion, is found in the intestinal brush border, where it cleaves the acidic propeptide from trypsinogen to yield active trypsin (Kitamoto, Y. et al. (1994) Proc. Natl. Acad. Sci. USA 91:7588-7592). Prolylcarboxypeptidase, a lysosomal serine peptidase that cleaves peptides such as angiotensin II and III and [des-Arg9] bradykinin, shares sequence homology with members of both the serine carboxypeptidase and prolylendopeptidase families (Tan, F. et al. (1993) J. Biol. Chem. 268:16631-16638). The protease neuropsin may influence synapse formation and neuronal connectivity in the hippocampus in response to neural signaling (Chen, Z.-L. et al. (1995) J. Neurosci. 15:5088-5097). Tissue plasminogen activator is useful for acute management of stroke (Zivin, J.A. (1999) Neurology 53:14-19) and myocardial infarction (Ross, A.M. (1999) Clin. Cardiol. 22:165-171). Some receptors (PAR, for proteinase-activated receptor), highly expressed throughout the digestive tract, are activated by proteolytic cleavage of an extracellular domain. The major agonists for PARs, thrombin, trypsin, and mast cell tryptase, are released in allergy and inflammatory conditions. Control of PAR activation by proteases has been suggested as a promising therapeutic target (Vergnolle, N. (2000) Aliment. Pharmacol. Ther. 14:257-266; Rice, K.D. et al. (1998) Curr. Pharm. Des. 4:381-396). Prostate-specific antigen (PSA) is a kallikrein-like serine protease synthesized and secreted exclusively by epithelial cells in the prostate gland. Serum PSA is elevated in prostate cancer and is the most sensitive physiological marker for monitoring cancer progression and response to therapy. PSA can also identify the prostate as the origin of a metastatic tumor (Brawer, M.K. and P.H. Lange (1989) Urology 33:11-16).

The signal peptidase is a specialized class of SP found in all prokaryotic and eukaryotic cell types that serves in the processing of signal peptides from certain proteins. Signal peptides are amino-terminal domains of a protein which direct the protein from its ribosomal assembly site to a particular cellular or extracellular location. Once the protein has been exported, removal of the signal sequence by a signal peptidase and posttranslational processing, e.g., glycosylation or phosphorylation, activate the protein. Signal peptidases exist as multi-subunit complexes in both yeast and mammals. The canine signal peptidase complex is composed of five subunits, all associated with the microsomal membrane and containing hydrophobic regions that span the membrane one or more times (Shelness, G.S. and G. Blobel (1990) J. Biol. Chem. 265:9512-9519). Some of these subunits serve to fix the complex in its proper position on the membrane while others contain the actual catalytic activity.

Thrombin is a serine protease with an essential role in the process of blood coagulation. Prothrombin, synthesized in the liver, is converted to active thrombin by Factor Xa. Activated thrombin then cleaves soluble fibrinogen to polymer-forming fibrin, a primary component of blood clots. In addition, thrombin activates Factor XIIIa, which plays a role in cross-linking fibrin.

Thrombin also stimulates platelet aggregation through proteolytic processing of a 41-residue amino-terminal peptide from protease-activated receptor 1 (PAR-1), formerly known as the thrombin receptor. The cleavage of the amino-terminal peptide exposes a new amino terminus and may also be associated with PAR-1 internalization (Stubbs, M.T. and W. Bode (1994) Curr. Opin. Struct. Biol. 4:823-832; and Ofoso, F.A. et al. (1998) Biochem. J. 336:283-285). In addition to stimulating platelet activation through cleavage of the PAR-1 receptor, thrombin also induces platelet aggregation following cleavage of glycoprotein V, also on the surface of platelets. Glycoprotein V appears to be the major thrombin substrate on intact platelets. Platelets deficient for glycoprotein V are hypersensitive to thrombin, which is still required to cleave PAR-1. While platelet aggregation is required for normal hemostasis in mammals, excessive platelet aggregation can result in arterial thrombosis, atherosclerotic arteries, acute myocardial infarction, and stroke (Ramakrishnan, V. et al. (1999) Proc. Natl. Acad. Sci. U.S.A. 96:13336-13341 and references within).

Proteases in another family have a serine in their active site and are dependent on the hydrolysis of ATP for their activity. These proteases contain proteolytic core domains and regulatory ATPase domains which can be identified by the presence of the P-loop, an ATP/GTP-binding motif (PROSITE PDOC00803). Members of this family include the eukaryotic mitochondrial matrix proteases, Clp protease and the proteasome. Clp protease was originally found in plant chloroplasts but is believed to be widespread in both prokaryotic and eukaryotic cells. The gene for early-onset torsion dystonia encodes a protein related to Clp protease (Ozelius, L.J. et al. (1998) Adv. Neurol. 78:93-105).

The proteasome is an intracellular protease complex found in some bacteria and in all eukaryotic cells, and plays an important role in cellular physiology. The proteasome is a large (~2000 kDa) multisubunit complex composed of a central catalytic core containing a variety of proteases arranged in four seven-membered rings with the active sites facing inwards into the central cavity, and terminal ATPase subunits covering the outer port of the cavity and regulating substrate entry (for review, see Schmidt, M. et al. (1999) Curr. Opin. Chem. Biol. 3:584-591). Proteasomes are associated with the ubiquitin conjugation system (UCS), a major pathway for the degradation of cellular proteins of all types, including proteins that function to activate or repress cellular processes such as transcription and cell cycle progression (Ciechanover, A. (1994) Cell 79:13-21). In the UCS pathway, proteins targeted for degradation are conjugated to ubiquitin, a small heat stable protein. The ubiquitinated protein is then recognized and degraded by the proteasome. The resultant ubiquitin-peptide complex is hydrolyzed by a ubiquitin carboxyl terminal hydrolase, and free ubiquitin is released for reutilization by the UCS. Ubiquitin-proteasome systems are implicated in the degradation of mitotic cyclic kinases, oncoproteins, tumor suppressor genes (p53), cell surface receptors associated with signal transduction, transcriptional regulators, and mutated or damaged

proteins (Ciechanover, *supra*). This pathway has been implicated in a number of diseases, including cystic fibrosis, Angelman's syndrome, and Liddle syndrome (reviewed in Schwartz, A.L. and A. Ciechanover (1999) Annu. Rev. Med. 50:57-74). A murine proto-oncogene, Unp, encodes a nuclear ubiquitin protease whose overexpression leads to oncogenic transformation of NIH3T3 cells. The human homolog of this gene is consistently elevated in small cell tumors and adenocarcinomas of the lung (Gray, D.A. (1995) Oncogene 10:2179-2183). Ubiquitin carboxyl terminal hydrolase is involved in the differentiation of a lymphoblastic leukemia cell line to a non-dividing mature state (Maki, A. et al. (1996) Differentiation 60:59-66). In neurons, ubiquitin carboxyl terminal hydrolase (PGP 9.5) expression is strong in the abnormal structures that occur in human neurodegenerative diseases (Lowe, J. et al. (1990) J. Pathol. 161:153-160).

Cysteine Proteases

Cysteine proteases (CPs) are involved in diverse cellular processes ranging from the processing of precursor proteins to intracellular degradation. Nearly half of the CPs known are present only in viruses. CPs have a cysteine as the major catalytic residue at the active site where catalysis proceeds via a thioester intermediate and is facilitated by nearby histidine and asparagine residues. A glutamine residue is also important, as it helps to form an oxyanion hole. Two important CP families include the papain-like enzymes (C1) and the calpains (C2). Papain-like family members are generally lysosomal or secreted and therefore are synthesized with signal peptides as well as propeptides. Most members bear a conserved motif in the propeptide that may have structural significance (Karrer, K.M. et al. (1993) Proc. Natl. Acad. Sci. USA 90:3063-3067). Three-dimensional structures of papain family members show a bilobed molecule with the catalytic site located between the two lobes. Papains include cathepsins B, C, H, L, and S, certain plant allergens and dipeptidyl peptidase (for a review, see Rawlings, N.D. and A.J. Barrett (1994) Methods Enzymol. 244:461-486).

Some CPs are expressed ubiquitously, while others are produced only by cells of the immune system. Of particular note, CPs are produced by monocytes, macrophages and other cells which migrate to sites of inflammation and secrete molecules involved in tissue repair. Overabundance of these repair molecules plays a role in certain disorders. In autoimmune diseases such as rheumatoid arthritis, secretion of the cysteine peptidase cathepsin C degrades collagen, laminin, elastin and other structural proteins found in the extracellular matrix of bones. Bone weakened by such degradation is also more susceptible to tumor invasion and metastasis. Cathepsin L expression may also contribute to the influx of mononuclear cells which exacerbates the destruction of the rheumatoid synovium (Keyszer, G.M. (1995) Arthritis Rheum. 38:976-984).

Calpains are calcium-dependent cytosolic endopeptidases which contain both an N-terminal catalytic domain and a C-terminal calcium-binding domain. Calpain is expressed as a proenzyme

heterodimer consisting of a catalytic subunit unique to each isoform and a regulatory subunit common to different isoforms. Each subunit bears a calcium-binding EF-hand domain. The regulatory subunit also contains a hydrophobic glycine-rich domain that allows the enzyme to associate with cell membranes. Calpains are activated by increased intracellular calcium concentration, which induces a change in conformation and limited autolysis. The resultant active molecule requires a lower calcium concentration for its activity (Chan, S.L. and M.P. Mattson (1999) J. Neurosci. Res. 58:167-190). Calpain expression is predominantly neuronal, although it is present in other tissues. Several chronic neurodegenerative disorders, including ALS, Parkinson's disease and Alzheimer's disease are associated with increased calpain expression (Chan and Mattson, *supra*). Calpain-mediated breakdown of the cytoskeleton has been proposed to contribute to brain damage resulting from head injury (McCracken, E. et al. (1999) J. Neurotrauma 16:749-761). Calpain-3 is predominantly expressed in skeletal muscle, and is responsible for limb-girdle muscular dystrophy type 2A (Minami, N. et al. (1999) J. Neurol. Sci. 171:31-37).

Another family of thiol proteases is the caspases, which are involved in the initiation and execution phases of apoptosis. A pro-apoptotic signal can activate initiator caspases that trigger a proteolytic caspase cascade, leading to the hydrolysis of target proteins and the classic apoptotic death of the cell. Two active site residues, a cysteine and a histidine, have been implicated in the catalytic mechanism. Caspases are among the most specific endopeptidases, cleaving after aspartate residues. Caspases are synthesized as inactive zymogens consisting of one large (p20) and one small (p10) subunit separated by a small spacer region, and a variable N-terminal prodomain. This prodomain interacts with cofactors that can positively or negatively affect apoptosis. An activating signal causes autoproteolytic cleavage of a specific aspartate residue (D297 in the caspase-1 numbering convention) and removal of the spacer and prodomain, leaving a p10/p20 heterodimer. Two of these heterodimers interact via their small subunits to form the catalytically active tetramer. The long prodomains of some caspase family members have been shown to promote dimerization and auto-processing of procaspases. Some caspases contain a "death effector domain" in their prodomain by which they can be recruited into self-activating complexes with other caspases and FADD protein associated death receptors or the TNF receptor complex. In addition, two dimers from different caspase family members can associate, changing the substrate specificity of the resultant tetramer. Endogenous caspase inhibitors (inhibitor of apoptosis proteins, or IAPs) also exist. All these interactions have clear effects on the control of apoptosis (reviewed in Chan and Mattson, *supra*; Salveson, G.S. and V.M. Dixit (1999) Proc. Natl. Acad. Sci. USA 96:10964-10967).

Caspases have been implicated in a number of diseases. Mice lacking some caspases have severe nervous system defects due to failed apoptosis in the neuroepithelium and suffer early lethality. Others show severe defects in the inflammatory response, as caspases are responsible for

processing IL-1b and possibly other inflammatory cytokines (Chan and Mattson, *supra*). Cowpox virus and baculoviruses target caspases to avoid the death of their host cell and promote successful infection. In addition, increases in inappropriate apoptosis have been reported in AIDS, neurodegenerative diseases and ischemic injury, while a decrease in cell death is associated with cancer (Salveson and Dixit, *supra*; Thompson, C.B. (1995) Science 267:1456-1462).

Aspartyl proteases

Aspartyl proteases (APs) include the lysosomal proteases cathepsins D and E, as well as chymosin, renin, and the gastric pepsins. Most retroviruses encode an AP, usually as part of the *pol* polyprotein. APs, also called acid proteases, are monomeric enzymes consisting of two domains, each domain containing one half of the active site with its own catalytic aspartic acid residue. APs are most active in the range of pH 2–3, at which one of the aspartate residues is ionized and the other neutral. The pepsin family of APs contains many secreted enzymes, and all are likely to be synthesized with signal peptides and propeptides. Most family members have three disulfide loops, the first ~5 residue loop following the first aspartate, the second 5-6 residue loop preceding the second aspartate, and the third and largest loop occurring toward the C terminus. Retropepsins, on the other hand, are analogous to a single domain of pepsin, and become active as homodimers with each retropepsin monomer contributing one half of the active site. Retropepsins are required for processing the viral polyproteins.

APs have roles in various tissues, and some have been associated with disease. Renin mediates the first step in processing the hormone angiotensin, which is responsible for regulating electrolyte balance and blood pressure (reviewed in Crews, D.E. and S.R. Williams (1999) Hum. Biol. 71:475-503). Abnormal regulation and expression of cathepsins are evident in various inflammatory disease states. Expression of cathepsin D is elevated in synovial tissues from patients with rheumatoid arthritis and osteoarthritis. The increased expression and differential regulation of the cathepsins are linked to the metastatic potential of a variety of cancers (Chambers, A.F. et al. (1993) Crit. Rev. Oncol. 4:95-114).

Metalloproteases

Metalloproteases require a metal ion for activity, usually manganese or zinc. Examples of manganese metalloenzymes include aminopeptidase P and human proline dipeptidase (PEPD). Aminopeptidase P can degrade bradykinin, a nonapeptide activated in a variety of inflammatory responses. Aminopeptidase P has been implicated in coronary ischemia/reperfusion injury. Administration of aminopeptidase P inhibitors has been shown to have a cardioprotective effect in rats (Ersahin, C. et al (1999) J. Cardiovasc. Pharmacol. 34:604-611).

Most zinc-dependent metalloproteases share a common sequence in the zinc-binding domain. The active site is made up of two histidines which act as zinc ligands and a catalytic glutamic acid C-

terminal to the first histidine. Proteins containing this signature sequence are known as the metzincins and include aminopeptidase N, angiotensin-converting enzyme, neurolysin, the matrix metalloproteases and the adamalysins (ADAMS). An alternate sequence is found in the zinc carboxypeptidases, in which all three conserved residues – two histidines and a glutamic acid – are involved in zinc binding.

A number of the neutral metalloendopeptidases, including angiotensin converting enzyme and the aminopeptidases, are involved in the metabolism of peptide hormones. High aminopeptidase B activity, for example, is found in the adrenal glands and neurohypophyses of hypertensive rats (Prieto, I. et al. (1998) Horm. Metab. Res. 30:246-248). Oligopeptidase M/neurolysin can hydrolyze bradykinin as well as neurotensin (Serizawa, A. et al. (1995) J. Biol. Chem 270:2092-2098). Neurotensin is a vasoactive peptide that can act as a neurotransmitter in the brain, where it has been implicated in limiting food intake (Tritos, N.A. et al. (1999) Neuropeptides 33:339-349).

The matrix metalloproteases (MMPs) are a family of at least 23 enzymes that can degrade components of the extracellular matrix (ECM). They are Zn^{2+} endopeptidases with an N-terminal catalytic domain. Nearly all members of the family have a hinge peptide and a C-terminal domain which can bind to substrate molecules in the ECM or to inhibitors produced by the tissue (TIMPs, for tissue inhibitor of metalloprotease; Campbell, I.L. and A. Pagenstecher (1999) Trends Neurosci. 22:285-287). The presence of fibronectin-like repeats, transmembrane domains, or C-terminal hemopexinase-like domains can be used to separate MMPs into collagenase, gelatinase, stromelysin and membrane-type MMP subfamilies. In the inactive form, the Zn^{2+} ion in the active site interacts with a cysteine in the pro-sequence. Activating factors disrupt the Zn^{2+}-cysteine interaction, or "cysteine switch," exposing the active site. This partially activates the enzyme, which then cleaves off its propeptide and becomes fully active. MMPs are often activated by the serine proteases plasmin and furin. MMPs are often regulated by stoichiometric, noncovalent interactions with inhibitors; the balance of protease to inhibitor, then, is very important in tissue homeostasis (reviewed in Yong, V.W. et al. (1998) Trends Neurosci. 21:75-80).

MMPs are implicated in a number of diseases including osteoarthritis (Mitchell, P. et al. (1996) J. Clin. Invest. 97:761-768), atherosclerotic plaque rupture (Sukhova, G.K. et al. (1999) Circulation 99:2503-2509), aortic aneurysm (Schneiderman, J. et al. (1998) Am. J. Path. 152:703-710), non-healing wounds (Saarialho-Kere, U.K. et al. (1994) J. Clin. Invest. 94:79-88), bone resorption (Blavier, L. and J.M. Delaisse (1995) J. Cell Sci. 108:3649-3659), age-related macular degeneration (Steen, B. et al. (1998) Invest. Ophthalmol. Vis. Sci. 39:2194-2200), emphysema (Finlay, G.A. et al. (1997) Thorax 52:502-506), myocardial infarction (Rohde, L.E. et al. (1999) Circulation 99:3063-3070) and dilated cardiomyopathy (Thomas, C.V. et al. (1998) Circulation 97:1708-1715). MMP inhibitors prevent metastasis of mammary carcinoma and experimental tumors

in rat, and Lewis lung carcinoma, hemangioma, and human ovarian carcinoma xenografts in mice (Eccles, S.A. et al. (1996) Cancer Res. 56:2815-2822; Anderson et al. (1996) Cancer Res. 56:715-718; Volpert, O.V. et al. (1996) J. Clin. Invest. 98:671-679; Taraboletti, G. et al. (1995) J. Natl. Cancer Inst. 87:293-298; Davies, B. et al. (1993) Cancer Res. 53:2087-2091). MMPs may be active in Alzheimer's disease. A number of MMPs are implicated in multiple sclerosis, and administration of MMP inhibitors can relieve some of its symptoms (reviewed in Yong et al., *supra*).

Another family of metalloproteases is the ADAMs, for A Disintegrin and Metalloprotease Domain, which they share with their close relatives the adamalysins, snake venom metalloproteases (SVMPs). ADAMs combine features of both cell surface adhesion molecules and proteases, containing a prodomain, a protease domain, a disintegrin domain, a cysteine rich domain, an epidermal growth factor repeat, a transmembrane domain, and a cytoplasmic tail. The first three domains listed above are also found in the SVMPs. The ADAMs possess four potential functions: proteolysis, adhesion, signaling and fusion. The ADAMs share the metzincin zinc binding sequence and are inhibited by some MMP antagonists such as TIMP-1.

ADAMs are implicated in such processes as sperm-egg binding and fusion, myoblast fusion, and protein-ectodomain processing or shedding of cytokines, cytokine receptors, adhesion proteins and other extracellular protein domains (Schlöndorff, J. and C.P. Blobel (1999) J. Cell. Sci. 112:3603-3617). The Kuzbanian protein cleaves a substrate in the NOTCH pathway (possibly NOTCH itself), activating the program for lateral inhibition in *Drosophila* neural development. Two ADAMs, TACE (ADAM 17) and ADAM 10, are proposed to have analogous roles in the processing of amyloid precursor protein in the brain (Schlöndorff and Blobel, *supra*). TACE has also been identified as the TNF activating enzyme (Black, R.A. et al. (1997) Nature 385:729-733). TNF is a pleiotropic cytokine that is important in mobilizing host defenses in response to infection or trauma, but can cause severe damage in excess and is often overproduced in autoimmune disease. TACE cleaves membrane-bound pro-TNF to release a soluble form. Other ADAMs may be involved in a similar type of processing of other membrane-bound molecules.

Proteins of the ADAMTS sub-family have all of the features of ADAM family metalloproteases and contain an additional thrombospondin domain (TS). The prototypic ADAMTS was identified in mouse, and found to be expressed in heart and kidney and upregulated by proinflammatory stimuli (Kuno, K. et al. (1997) J. Biol. Chem. 272:556-562). To date eleven members are recognized by the Human Genome Organization (HUGO; http://www.gene.ucl.ac.uk/users/hester/adamts.html#Approved). Members of this family have the ability to degrade aggrecan, a high molecular weight proteoglycan which provides cartilage with important mechanical properties including compressibility, and which is lost during the development of arthritis. Enzymes which degrade aggrecan are thus considered attractive targets to prevent and

slow the degradation of articular cartilage (See, e.g., Tortorella, M.D. (1999) Science 284:1664-1666; Abbaszade, I. (1999) J. Biol. Chem. 274:23443-23450). Other members are reported to have antiangiogenic potential (Kuno et al., *supra*) and/or procollagen processing (Colige, A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2374-2379).

Protease inhibitors

Protease inhibitors and other regulators of protease activity control the activity and effects of proteases. Protease inhibitors have been shown to control pathogenesis in animal models of proteolytic disorders (Murphy, G. (1991) Agents Actions Suppl. 35:69-76). Low levels of the cystatins, low molecular weight inhibitors of the cysteine proteases, correlate with malignant progression of tumors (Calkins, C. et al. (1995) Biol. Biochem. Hoppe Seyler 376:71-80). The cystatin superfamily of protease inhibitors is characterized by a particular pattern of linearly arranged and tandemly repeated disulfide loops (Kellermann, J. et al. (1989) J. Biol. Chem. 264:14121-14128). Serpins are inhibitors of mammalian plasma serine proteases. Many serpins serve to regulate the blood clotting cascade and/or the complement cascade in mammals. Sp32 is a positive regulator of the mammalian acrosomal protease, acrosin, that binds the proenzyme, proacrosin, and thereby aides in packaging the enzyme into the acrosomal matrix (Baba, T. et al. (1994) J. Biol. Chem. 269:10133-10140). The Kunitz family of serine protease inhibitors are characterized by one or more "Kunitz domains" containing a series of cysteine residues that are regularly spaced over approximately 50 amino acid residues and form three intrachain disulfide bonds. Members of this family include aprotinin, tissue factor pathway inhibitor (TFPI-1 and TFPI-2), inter-a-trypsin inhibitor, and bikunin (Marlor, C.W. et al. (1997) J. Biol. Chem. 272:12202-12208). Members of this family are potent inhibitors (in the nanomolar range) against serine proteases such as kallikrein and plasmin. Aprotinin has clinical utility in reduction of perioperative blood loss.

A major portion of all proteins synthesized in eukaryotic cells are synthesized on the cytosolic surface of the endoplasmic reticulum (ER). Before these immature proteins are distributed to other organelles in the cell or are secreted, they must be transported into the interior lumen of the ER where post-translational modifications are performed. These modifications include protein folding and the formation of disulfide bonds, and N-linked glycosylations.

Protein Isomerases

Protein folding in the ER is aided by two principal types of protein isomerases, protein disulfide isomerase (PDI), and peptidyl-prolyl isomerase (PPI). PDI catalyzes the oxidation of free sulfhydryl groups in cysteine residues to form intramolecular disulfide bonds in proteins. PPI, an enzyme that catalyzes the isomerization of certain proline imidic bonds in oligopeptides and proteins, is considered to govern one of the rate limiting steps in the folding of many proteins to their final functional conformation. The cyclophilins represent a major class of PPI that was originally

identified as the major receptor for the immunosuppressive drug cyclosporin A (Handschumacher, R.E. et al. (1984) Science 226: 544-547).

Protein Glycosylation

The glycosylation of most soluble secreted and membrane-bound proteins by oligosaccharides linked to asparagine residues in proteins is also performed in the ER. This reaction is catalyzed by a membrane-bound enzyme, oligosaccharyl transferase. Although the exact purpose of this "N-linked" glycosylation is unknown, the presence of oligosaccharides tends to make a glycoprotein resistant to protease digestion. In addition, oligosaccharides attached to cell-surface proteins called selectins are known to function in cell-cell adhesion processes (Alberts, B. et al. (1994) Molecular Biology of the Cell Garland Publishing Co., New York, NY, p. 608). "O-linked" glycosylation of proteins also occurs in the ER by the addition of N-acetylgalactosamine to the hydroxyl group of a serine or threonine residue followed by the sequential addition of other sugar residues to the first. This process is catalyzed by a series of glycosyltransferases, each specific for a particular donor sugar nucleotide and acceptor molecule (Lodish, H. et al. (1995) Molecular Cell Biology, W. H. Freeman and Co., New York, NY, pp. 700-708). In many cases, both N- and O-linked oligosaccharides appear to be required for the secretion of proteins or the movement of plasma membrane glycoproteins to the cell surface.

An additional glycosylation mechanism operates in the ER specifically to target lysosomal enzymes to lysosomes and prevent their secretion. Lysosomal enzymes in the ER receive an N-linked oligosaccharide, like plasma membrane and secreted proteins, but are then phosphorylated on one or two mannose residues. The phosphorylation of mannose residues occurs in two steps, the first step being the addition of an N-acetylglucosamine phosphate residue by N-acetylglucosamine phosphotransferase, and the second the removal of the N-acetylglucosamine group by phosphodiesterase. The phosphorylated mannose residue then targets the lysosomal enzyme to a mannose 6-phosphate receptor which transports it to a lysosome vesicle (Lodish et al. *supra*, pp. 708-711).

Chaperones

Molecular chaperones are proteins that aid in the proper folding of immature proteins and refolding of improperly folded ones, the assembly of protein subunits, and in the transport of unfolded proteins across membranes. Chaperones are also called heat-shock proteins (hsp) because of their tendency to be expressed in dramatically increased amounts following brief exposure of cells to elevated temperatures. This latter property most likely reflects their need in the refolding of proteins that have become denatured by the high temperatures. Chaperones may be divided into several classes according to their location, function, and molecular weight, and include hsp60, TCP1, hsp70, hsp40 (also called DnaJ), and hsp90. For example, hsp90 binds to steroid hormone receptors,

represses transcription in the absence of the ligand, and provides proper folding of the ligand-binding domain of the receptor in the presence of the hormone (Burston, S.G. and A.R. Clarke (1995) Essays Biochem. 29:125-136). Hsp60 and hsp70 chaperones aid in the transport and folding of newly synthesized proteins. Hsp70 acts early in protein folding, binding a newly synthesized protein before it leaves the ribosome and transporting the protein to the mitochondria or ER before releasing the folded protein. Hsp60, along with hsp10, binds misfolded proteins and gives them the opportunity to refold correctly. All chaperones share an affinity for hydrophobic patches on incompletely folded proteins and the ability to hydrolyze ATP. The energy of ATP hydrolysis is used to release the hsp-bound protein in its properly folded state (Alberts et al., *supra*, pp. 214, 571-572).

Expression profiling

Microarrays are analytical tools used in bioanalysis. A microarray has a plurality of molecules spatially distributed over, and stably associated with, the surface of a solid support. Microarrays of polypeptides, polynucleotides, and/or antibodies have been developed and find use in a variety of applications, such as gene sequencing, monitoring gene expression, gene mapping, bacterial identification, drug discovery, and combinatorial chemistry.

One area in particular in which microarrays find use is in gene expression analysis. Array technology can provide a simple way to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling cascade, carry out housekeeping functions, or are specifically related to a particular genetic predisposition, condition, disease, or disorder.

Lung cancer is the leading cause of cancer death in the United States, affecting more than 100,000 men and 50,000 women each year. Nearly 90% of the patients diagnosed with lung cancer are cigarette smokers. Tobacco smoke contains thousands of noxious substances that induce carcinogen metabolizing enzymes and covalent DNA adduct formation in the exposed bronchial epithelium. In nearly 80% of patients diagnosed with lung cancer, metastasis has already occurred. Most commonly lung cancers metastasize to pleura, brain, bone, pericardium, and liver. This adversely affects the overall five-year survival rate which is 37% for squamous carcinoma, 27% for adenocarcinoma and large cell carcinoma, and less than 1% for small cell carcinomas. Earlier diagnosis and an systematic approach to identification, staging, and treatment could positively affect patient outcome (DeVita et al. (1997) Cancer: Principles and Practice of Oncology, Lippincott-Raven, Philadelphia PA) and Fauci et al. (1998) Harrison's Principals of Internal Medicine, McGraw Hill, New York, NY).

Lung cancers progress through a series of morphologically distinct stages from hyperplasia to invasive carcinoma. Malignant lung cancers are divided into two groups comprising four histopathological classes. The nonsmall cell lung carcinoma (NSCLC) group includes squamous cell carcinomas, adenocarcinomas, and large cell carcinomas and accounts for about 70% of all lung cancer cases. Adenocarcinomas typically arise in the peripheral airways and often form mucin secreting glands. Squamous cell carcinomas typically arise in proximal airways. The histogenesis of squamous cell carcinomas may be related to chronic inflammation and injury to the bronchial epithelium, leading to squamous metaplasia. The small cell lung carcinoma (SCLC) group accounts for about 20% of lung cancer cases. SCLCs typically arise in proximal airways and exhibit a number of paraneoplastic syndromes including inappropriate production of adrenocorticotropin and anti-diuretic hormone.

Lung cancer cells accumulate numerous genetic lesions, many of which are associated with cytologically visible chromosomal aberrations. The high frequency of chromosomal deletions associated with lung cancer may reflect the role of multiple tumor suppressor loci in the etiology of this disease. Several studies report deletions of regions of chromosome 11 in NSCLC (Bepler, G. and Garcia-Blanco, M.A. (1994) Proc. Natl. Acad. Sci. USA 91:5513-7; Iizuka, M., et al. (1995) Genes, Chromosomes & Cancer 13:40-46; Rasio, D. (1995) Cancer Research 55:3988-91). Deletions in other chromosome arms such as 3p, 9p and 17p are also common. Other frequently observed genetic lesions include overexpression of telomerase, activation of oncogenes such as K-ras and c-myc, and inactivation of tumor suppressor genes such as RB, p53 and p16 (Toomey, D. et al.(2001) Cancer 92:2648-57; Zajac-Kaye M. (2001) Lung Cancer 34:S43-6; Wright, G. et al. (2000) Current Opinion in Oncology 12:143-8; Kohno, T. and Yokota, J. (1999) Carcinogenesis 20:1403-10).

Prostate cancer is a common malignancy in men over the age of 50, and the incidence increases with age. In the US, there are approximately 132,000 newly diagnosed cases of prostate cancer and more than 33,000 deaths from the disorder each year.

Once cancer cells arise in the prostate, they are stimulated by testosterone to a more rapid growth. Thus, removal of the testes can indirectly reduce both rapid growth and metastasis of the cancer. Over 95 percent of prostatic cancers are adenocarcinomas which originate in the prostatic acini. The remaining 5 percent are divided between squamous cell and transitional cell carcinomas, both of which arise in the prostatic ducts or other parts of the prostate gland.

As with most cancers, prostate cancer develops through a multistage progression ultimately resulting in an aggressive, metastatic phenotype. The initial step in tumor progression involves the hyperproliferation of normal luminal and/or basal epithelial cells that become hyperplastic and evolve

into early-stage tumors. The early-stage tumors are localized in the prostate but eventually may metastasize, particularly to the bone, brain or lung. About 80% of these tumors remain responsive to androgen treatment, an important hormone controlling the growth of prostate epithelial cells. However, in its most advanced state, cancer growth becomes androgen-independent and there is currently no known treatment for this condition.

A primary diagnostic marker for prostate cancer is prostate specific antigen (PSA). PSA is a tissue-specific serine protease almost exclusively produced by prostatic epithelial cells. The quantity of PSA correlates with the number and volume of the prostatic epithelial cells, and consequently, the levels of PSA are an excellent indicator of abnormal prostate growth. Men with prostate cancer exhibit an early linear increase in PSA levels followed by an exponential increase prior to diagnosis. However, since PSA levels are also influenced by factors such as inflammation, androgen and other growth factors, some scientists maintain that changes in PSA levels are not useful in detecting individual cases of prostate cancer.

Current areas of cancer research provide additional prospects for markers as well as potential therapeutic targets for prostate cancer. Several growth factors have been shown to play a critical role in tumor development, growth, and progression. The growth factors Epidermal Growth Factor (EGF), Fibroblast Growth Factor (FGF), and Tumor Growth Factor alpha (TGFα) are important in the growth of normal as well as hyperproliferative prostate epithelial cells, particularly at early stages of tumor development and progression, and affect signaling pathways in these cells in various ways (Lin J et al. (1999) Cancer Res. 59:2891-2897; Putz T et al. (1999) Cancer Res 59:227-233). The TGF-β family of growth factors are generally expressed at increased levels in human cancers and the high expression levels in many cases correlates with advanced stages of malignancy and poor survival (Gold LI (1999) Crit Rev Oncog 10:303-360). Finally, there are human cell lines representing both the androgen-dependent stage of prostate cancer (LNCap) as well as the androgen-independent, hormone refractory stage of the disease (PC3 and DU-145) that have proved useful in studying gene expression patterns associated with the progression of prostate cancer, and the effects of cell treatments on these expressed genes (Chung TD (1999) Prostate 15:199-207).

Ovarian cancer is the leading cause of death from a gynecologic cancer. The majority of ovarian cancers are derived from epithelial cells, and 70% of patients with epithelial ovarian cancers present with late-stage disease. As a result the loingterm survival rates for this disease are very low. Identification of early stage markers for ovarian cancer would significantly increase the survival rate. The molecular events that lead to ovarian cancer are poorly understood. Some of the known aberrations include mutation of p53 and microsatellite instability. Osteosarcoma is the most common malignant bone tumor in children. With currently available treatment regimens, approximately 30–40% of patients with non-metastatic disease relapse after therapy. Currently, there is no

prognostic factor that can be used at the time of initial diagnosis to predict which patients will have a high risk of relapse. The only significant prognostic factor predicting the outcome in a patient with non-metastatic osteosarcoma is the histopathologic response of the primary tumor resected at the time of definitive surgery.

There are more than 180,000 new cases of breast cancer diagnosed each year, and the mortality rate for breast cancer approaches 10% of all deaths in females between the ages of 45-54 (K. Gish (1999) AWIS Magazine 28:7-10). However the survival rate based on early diagnosis of localized breast cancer is extremely high (97%), compared with the advanced stage of the disease in which the tumor has spread beyond the breast (22%). Current procedures for clinical breast examination are lacking in sensitivity and specificity, and efforts are underway to develop comprehensive gene expression profiles for breast cancer that may be used in conjunction with conventional screening methods to improve diagnosis and prognosis of this disease (Perou CM et al. (2000) Nature 406:747-752).

Breast cancer is a genetic disease commonly caused by mutations in cellular disease. Mutations in two genes, BRCA1 and BRCA2, are known to greatly predispose a woman to breast cancer and may be passed on from parents to children (Gish, supra). However, this type of hereditary breast cancer accounts for only about 5% to 9% of breast cancers, while the vast majority of breast cancer is due to noninherited mutations that occur in breast epithelial cells.

A good deal is already known about the expression of specific genes associated with breast cancer. For example, the relationship between expression of epidermal growth factor (EGF) and its receptor, EGFR, to human mammary carcinoma has been particularly well studied. (See Khazaie et al., supra, and references cited therein for a review of this area.) Overexpression of EGFR, particularly coupled with down-regulation of the estrogen receptor, is a marker of poor prognosis in breast cancer patients. In addition, EGFR expression in breast tumor metastases is frequently elevated relative to the primary tumor, suggesting that EGFR is involved in tumor progression and metastasis. This is supported by accumulating evidence that EGF has effects on cell functions related to metastatic potential, such as cell motility, chemotaxis, secretion and differentiation. Changes in expression of other members of the erbB receptor family, of which EGFR is one, have also been implicated in breast cancer. The abundance of erbB receptors, such as HER-2/neu, HER-3, and HER-4, and their ligands in breast cancer points to their functional importance in the pathogenesis of the disease, and may therefore provide targets for therapy of the disease (Bacus, SS et al. (1994) Am J Clin Pathol 102:S13-S24). Other known markers of breast cancer include a human secreted frizzled protein mRNA that is downregulated in breast tumors; the matrix G1a protein which is overexpressed is human breast carcinoma cells; Drg1 or RTP, a gene whose expression is diminished in colon, breast, and prostate tumors; maspin, a tumor suppressor gene downregulated in invasive breast

carcinomas; and CaN19, a member of the S100 protein family, all of which are down regulated in mammary carcinoma cells relative to normal mammary epithelial cells (Zhou Z et al. (1998) Int J Cancer 78:95-99; Chen, L et al. (1990) Oncogene 5:1391-1395; Ulrix W et al (1999) FEBS Lett 455:23-26; Sager, R et al. (1996) Curr Top Microbiol Immunol 213:51-64; and Lee, SW et al. (1992) Proc Natl Acad Sci USA 89:2504-2508).

Cell lines derived from human mammary epithelial cells at various stages of breast cancer provide a useful model to study the process of malignant transformation and tumor progression as it has been shown that these cell lines retain many of the properties of their parental tumors for lengthy culture periods (Wistuba II et al. (1998) Clin Cancer Res 4:2931-2938). Such a model is particularly useful for comparing phenotypic and molecular characteristics of human mammary epithelial cells at various stages of malignant transformation.

Colon cancer evolves through a multi-step process whereby pre-malignant colonocytes undergo a relatively defined sequence of events leading to tumor formation. While soft tissue sarcomas are relatively rare, more than 50% of new patients diagnosed with the disease will die from it. The molecular pathways leading to the development of sarcomas are relatively unknown, due to the rarity of the disease and variation in pathology. Several factors participate in the process of tumor progression and malignant transformation including genetic factors, mutations, and selection.

To understand the nature of gene alterations in colorectal cancer, a number of studies have focused on the inherited syndromes. Familial adenomatous polyposis (FAP), is caused by mutations in the adenomatous polyposis coli gene (APC), resulting in truncated or inactive forms of the protein. This tumor suppressor gene has been mapped to chromosome 5q. Hereditary nonpolyposis colorectal cancer (HNPCC) is caused by mutations in mis-match repair genes. Although hereditary colon cancer syndromes occur in a small percentage of the population and most colorectal cancers are considered sporadic, knowledge from studies of the hereditary syndromes can be generally applied. For instance, somatic mutations in APC occur in at least 80% of sporadic colon tumors. APC mutations are thought to be the initiating event in the disease. Other mutations occur subsequently. Approximately 50% of colorectal cancers contain activating mutations in ras, while 85% contain inactivating mutations in p53. Changes in all of these genes lead to gene expression changes in colon cancer.

Steroids are a class of lipid-soluble molecules, including cholesterol, bile acids, vitamin D, and hormones, that share a common four-ring structure based on cyclopentano-perhydrophenanthrene and that carry out a wide variety of functions. Cholesterol, for example, is a component of cell membranes that controls membrane fluidity. It is also a precursor for bile acids which solubilize lipids and facilitate absorption in the small intestine during digestion. Vitamin D regulates the absorption of calcium in the small intestine and controls the concentration of calcium in plasma. Steroid hormones, produced by the adrenal cortex, ovaries, and testes, include glucocorticoids,

mineralocorticoids, androgens, and estrogens. They control various biological processes by binding to intracellular receptors that regulate transcription of specific genes in the nucleus. Glucocorticoids, for example, increase blood glucose concentrations by regulation of gluconeogenesis in the liver, increase blood concentrations of fatty acids by promoting lipolysis in adipose tissues, modulate sensitivity to catcholamines in the central nervous system, and reduce inflammation. The principal mineralocorticoid, aldosterone, is produced by the adrenal cortex and acts on cells of the distal tubules of the kidney to enhance sodium ion reabsorption. Androgens, produced by the interstitial cells of Leydig in the testis, include the male sex hormone testosterone, which triggers changes at puberty, the production of sperm and maintenance of secondary sexual characteristics. Female sex hormones, estrogen and progesterone, are produced by the ovaries and also by the placenta and adrenal cortex of the fetus during pregnancy. Estrogen regulates female reproductive processes and secondary sexual characteristics. Progesterone regulates changes in the endometrium during the menstrual cycle and pregnancy.

Steroid hormones are widely used for fertility control and in anti-inflammatory treatments for physical injuries and diseases such as arthritis, asthma, and auto-immune disorders. Progesterone, a naturally occurring progestin, is primarily used to treat amenorrhea, abnormal uterine bleeding, or as a contraceptive. Endogenous progesterone is responsible for inducing secretory activity in the endometrium of the estrogen-primed uterus in preparation for the implantation of a fertilized egg and for the maintenance of pregnancy. It is secreted from the corpus luteum in response to luteinizing hormone (LH). The primary contraceptive effect of exogenous progestins involves the suppression of the midcycle surge of LH. At the cellular level, progestins diffuse freely into target cells and bind to the progesterone receptor. Target cells include the female reproductive tract, the mammary gland, the hypothalamus, and the pituitary. Once bound to the receptor, progestins slow the frequency of release of gonadotropin releasing hormone from the hypothalamus and blunt the pre-ovulatory LH surge, thereby preventing follicular maturation and ovulation. Progesterone has minimal estrogenic and androgenic activity. Progesterone is metabolized hepatically to pregnanediol and conjugated with glucuronic acid.

Corticosteroids are used to relieve inflammation and to suppress the immune response. They inhibit eosinophil, basophil, and airway epithelial cell function by regulation of cytokines that mediate the inflammatory response. They inhibit leukocyte infiltration at the site of inflammation, interfere in the function of mediators of the inflammatory response, and suppress the humoral immune response. Corticosteroids are used to treat allergies, asthma, arthritis, and skin conditions. Budesonide is a corticosteroid used to control symptoms associated with allergic rhinitis or asthma. Budesonide has high topical anti-inflammatory activity but low systemic activity.

The most important function of adipose tissue is its ability to store and release fat during periods of feeding and fasting. White adipose tissue is the major energy reserve in periods of excess energy use. Its primary purpose is mobilization during energy deprivation. Understanding how various molecules regulate adiposity and energy balance in physiological and pathophysiological situations may lead to the development of novel therapeutics for human obesity. Adipose tissue is also one of the important target tissues for insulin. Adipogenesis and insulin resistance in type II diabetes are linked and present intriguing relations. Most patients with type II diabetes are obese and obesity in turn causes insulin resistance.

The majority of research in adipocyte biology to date has been done using transformed mouse preadipocyte cell lines. The culture condition which stimulates mouse preadipocyte differentiation is different from that for inducing human primary preadipocyte differentiation. In addition, primary cells are diploid and may therefore reflect the <u>in vivo</u> context better than aneuploid cell lines. Understanding the gene expression profile during adipogenesis in humans will lead to understanding the fundamental mechanism of adiposity regulation. Furthermore, through comparing the gene expression profiles of adipogenesis between donor with normal weight and donor with obesity, identification of crucial genes, potential drug targets for obesity and type II diabetes, will be possible.

Insulin sensitivity can be enahnced by various compounds. Thiazolidinediones (TZDs) act as agonists for the peroxisome-proliferator-activated receptor gamma (PPARγ), a member of the nuclear hormone receptor superfamily. TZDs reduce hyperglycemia, hyperinsulinemia, and hypertension, in part by promoting glucose metabolism and inhibiting gluconeogenesis. Roles for PPARγ and its agonists have been demonstrated in a wide range of pathological conditions including diabetes, obesity, hypertension, atherosclerosis, polycystic ovarian syndrome, and cancers such as breast, prostate, liposarcoma, and colon cancer.

The mechanism by which TZDs and other PPARγ agonists enhance insulin sensitivity is not fully understood, but may involve the ability of PPARγ to promote adipogenesis. When ectopically expressed in cultured preadipocytes, PPARγ is a potent inducer of adipocyte differentiation. TZDs, in combination with insulin and other factors, can also enhance differentiation of human preadipocytes in culture (Adams et al. (1997) J. Clin. Invest. 100:3149-3153). The relative potency of different TZDs in promoting adipogenesis <u>in vitro</u> is proportional to both their insulin sensitizing effects <u>in vivo</u>, and their ability to bind and activate PPARγ <u>in vitro</u>. Interestingly, adipocytes derived from omental adipose depots are refractory to the effects of TZDs. It has therefore been suggested that the insulin sensitizing effects of TZDs may result from their ability to promote adipogenesis in subcutaneous adipose depots (Adams et al., <u>supra</u>). Further, dominant negative mutations in the PPARγ gene have been identified in two non-obese subjects with severe insulin resistance,

hypertension, and overt non-insulin dependent diabetes mellitus (NIDDM) (Barroso et al. (1998) Nature 402:880-883).

NIDDM is the most common form of diabetes mellitus, a chronic metabolic disease that affects 143 million people worldwide. NIDDM is characterized by abnormal glucose and lipid metabolism that result from a combination of peripheral insulin resistance and defective insulin secretion. NIDDM has a complex, progressive etiology and a high degree of heritability. Numerous complications of diabetes including heart disease, stroke, renal failure, retinopathy, and peripheral neuropathy contribute to the high rate of morbidity and mortality.

Tangier disease (TD) is a genetic disorder characterized by near absence of circulating high density lipoprotein (HDL) and the accumulation of cholesterol esters in many tissues, including tonsils, lymph nodes, liver, spleen, thymus, and intestine. Low levels of HDL represent a clear predictor of premature coronary artery disease and homozygous TD correlates with a four- to six-fold increase in cardiovascular disease compared to controls. HDL plays a cardio-protective role in reverse cholesterol transport, the flux of cholesterol from peripheral cells such as tissue macrophages through plasma lipoproteins to the liver. The HDL protein, apolipoprotein A-I, plays a major role in this process, interacting with the cell surface to remove excess cholesterol and phospholipids. This pathway is severely impaired in TD and the defect lies in a specific gene, the ABC1 transporter. This gene is a member of the family of ATP-binding cassette transporters, which utilize ATP hydrolysis to transport a variety of substrates across membranes.

Leukocytes comprise lymphocytes, granulocytes, and monocytes. Lymphocytes include T- and B-cells, which specifically recognize and respond to foreign pathogens. T-cells fight viral infections and activate other leukocytes, while B-cells secrete antibodies that neutralize bacteria and other microbes. Granulocytes and monocytes are primarily migratory, phagocytic cells that exit the bloodstream to fight infection in tissues. Monocytes, which are derived from immature promonocytes, further differentiate into macrophages that engulf and digest microorganisms and damaged or dead cells. Monocytes and macrophages modulate the immune response by secreting signaling molecules such as growth factors and cytokines. Tumor necrosis factor-α (TNF-α), for example, is a macrophage-secreted protein with anti-tumor and anti-viral activity. In addition, monocytes and macrophages are recruited to sites of infection and inflammation by signaling proteins secreted by other leukocytes. The differentiation of the monocyte blood cell lineage can be studied in vitro using cultured cell lines. For example, THP-1 is a human promonocyte cell line that can be activated by treatment with both phorbol ester such as phorbol myristate acetate (PMA), and lipopolysaccharide (LPS). PMA is a broad activator of the protein kinase C-dependent pathways.

Monocytes are involved in the initiation and maintenance of inflammatory immune responses. The outer membrane of gram-negative bacteria expresses lipopolysaccharide (LPS)

complexes called endotoxins. Toxicity is associated with the lipid component (Lipid A) of LPS, and immunogenicity is associated with the polysaccharide components of LPS. LPS elicits a variety of inflammatory responses, and because it activates complement by the alternative (properdin) pathway, it is often part of the pathology of gram-negative bacterial infections. For the most part, endotoxins remain associated with the cell wall until the bacteria disintegrate. LPS released into the bloodstream by lysing gram-negative bacteria is first bound by certain plasma proteins identified as LPS-binding proteins. The LPS-binding protein complex interacts with CD14 receptors on monocytes, macrophages, B cells, and other types of receptors on endothelial cells. Activation of human B cells with LPS results in mitogenesis as well as immunoglobulin synthesis. In monocytes and macrophages three types of events are triggered during their interaction with LPS: 1) Production of cytokines, including IL-1, IL-6, IL-8, TNF-α, and platelet-activating factor, which stimulate production of prostaglandins and leukotrienes that mediate inflammation and septic shock; 2) Activation of the complement cascade; and 3) Activation of the coagulation cascade.

Osteoarthritis (OA) is a debilitating joint disease involving focal cartilage loss. Several studies indicate a major genetic component can be involved in causing OA. Estimates of inheritability from twin studies of radiographic OA of the hand, knee and hip range from 36% to 68% (MacGregor, A.J. and Spector, T.D. (1999) Rheumatology 38:583-560). Several interleukin and interleukin-associated genes are located at 2q12-q22 (Leppavouri, J. et al. (1999) Am. J. Hum. Genet. 65:1060-1067). Interleukins regulate a number of enzymes that degrade the cartilage extracellular matrix, and the expression of certain interleukin genes, including IL-1β, is altered in OA joint tissue (Elson, C.J. et al. (1998) Br. J. Rheum. 37:106-107.

There is a need in the art for new compositions, including nucleic acids and proteins, for the diagnosis, prevention, and treatment of gastrointestinal, cardiovascular, autoimmune/inflammatory, cell proliferative, developmental, epithelial, neurological, and reproductive disorders.

SUMMARY OF THE INVENTION

Various embodiments of the invention provide purified polypeptides, protein modification and maintenance molecules, referred to collectively as 'PMMM' and individually as 'PMMM-1,' 'PMMM-2,' 'PMMM-3,' 'PMMM-4,' 'PMMM-5,' 'PMMM-6,' 'PMMM-7,' 'PMMM-8,' 'PMMM-9,' 'PMMM-10,' 'PMMM-11,' 'PMMM-12,' 'PMMM-13,' 'PMMM-14,' 'PMMM-15,' 'PMMM-16,' 'PMMM-17,' 'PMMM-18,' 'PMMM-19,' 'PMMM-20,' 'PMMM-21,' 'PMMM-22,' 'PMMM-23,' 'PMMM-24,' 'PMMM-25,' 'PMMM-26,' 'PMMM-27,' 'PMMM-28,' 'PMMM-29,' 'PMMM-30,' 'PMMM-31,' 'PMMM-32,' 'PMMM-33,' 'PMMM-34,' 'PMMM-35,' 'PMMM-36,' 'PMMM-37,' 'PMMM-38,' 'PMMM-39,' 'PMMM-40,' 'PMMM-41,' 'PMMM-42,' 'PMMM-43,' 'PMMM-44,' 'PMMM-45,' 'PMMM-46,' 'PMMM-47,' 'PMMM-48,' 'PMMM-49,' 'PMMM-50,' 'PMMM-

51,' 'PMMM-52,' 'PMMM-53,' 'PMMM-54,' 'PMMM-55,' 'PMMM-56,' 'PMMM-57,' and 'PMMM-58' and methods for using these proteins and their encoding polynucleotides for the detection, diagnosis, and treatment of diseases and medical conditions. Embodiments also provide methods for utilizing the purified protein modification and maintenance molecules and/or their encoding polynucleotides for facilitating the drug discovery process, including determination of efficacy, dosage, toxicity, and pharmacology. Related embodiments provide methods for utilizing the purified protein modification and maintenance molecules and/or their encoding polynucleotides for investigating the pathogenesis of diseases and medical conditions.

An embodiment provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58. Another embodiment provides an isolated polypeptide comprising an amino acid sequence of SEQ ID NO:1-58.

Still another embodiment provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58. In another embodiment, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-58. In an alternative embodiment, the polynucleotide is selected from the group consisting of SEQ ID NO:59-116.

Still another embodiment provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ

ID NO:1-58. Another embodiment provides a cell transformed with the recombinant polynucleotide. Yet another embodiment provides a transgenic organism comprising the recombinant polynucleotide.

Another embodiment provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Yet another embodiment provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58.

Still yet another embodiment provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In other embodiments, the polynucleotide can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID

NO:59-116, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex. In a related embodiment, the method can include detecting the amount of the hybridization complex. In still other embodiments, the probe can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Still yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof. In a related embodiment, the method can include detecting the amount of the amplified target polynucleotide or fragment thereof.

Another embodiment provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and a pharmaceutically acceptable excipient. In one embodiment, the composition can comprise an amino acid sequence selected from the group consisting of SEQ ID NO:1-58. Other embodiments provide a method of treating a disease or condition associated with decreased or abnormal expression of functional PMMM, comprising administering to a patient in need of such treatment the composition.

Yet another embodiment provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino

acid sequence selected from the group consisting of SEQ ID NO:1-58, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. Another embodiment provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with decreased expression of functional PMMM, comprising administering to a patient in need of such treatment the composition.

Still yet another embodiment provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. Another embodiment provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with overexpression of functional PMMM, comprising administering to a patient in need of such treatment the composition.

Another embodiment provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the

polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

Yet another embodiment provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

Still yet another embodiment provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

Another embodiment provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting

of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide can comprise a fragment of a polynucleotide selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for full length polynucleotide and polypeptide embodiments of the invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog, and the PROTEOME database identification numbers and annotations of PROTEOME database homologs, for polypeptide embodiments of the invention. The probability scores for the matches between each polypeptide and its homolog(s) are also shown.

Table 3 shows structural features of polypeptide embodiments, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide embodiments, along with selected fragments of the polynucleotides.

Table 5 shows representative cDNA libraries for polynucleotide embodiments.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze polynucleotides and polypeptides, along with applicable descriptions, references, and threshold parameters.

Table 8 shows single nucleotide polymorphisms found in polynucleotide sequences of the invention, along with allele frequencies in different human populations.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleic acids, and methods are described, it is understood that embodiments of the invention are not limited to the particular machines, instruments, materials, and

methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the invention.

As used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with various embodiments of the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"PMMM" refers to the amino acid sequences of substantially purified PMMM obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of PMMM. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of PMMM either by directly interacting with PMMM or by acting on components of the biological pathway in which PMMM participates.

An "allelic variant" is an alternative form of the gene encoding PMMM. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding PMMM include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as PMMM or a polypeptide with at least one functional characteristic of PMMM. Included within this definition

are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding PMMM, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide encoding PMMM. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent PMMM. Deliberate amino acid substitutions may be made on the basis of one or more similarities in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of PMMM is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" can refer to an oligopeptide, a peptide, a polypeptide, or a protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid. Amplification may be carried out using polymerase chain reaction (PCR) technologies or other nucleic acid amplification technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of PMMM. Antagonists may include proteins such as antibodies, anticalins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of PMMM either by directly interacting with PMMM or by acting on components of the biological pathway in which PMMM participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind PMMM polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin,

thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an *in vitro* evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH_2), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker (Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13).

The term "intramer" refers to an aptamer which is expressed *in vivo*. For example, a vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl. Acad. Sci. USA 96:3606-3610).

The term "spiegelmer" refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on substrates containing right-handed nucleotides.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a polynucleotide having a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis

or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic PMMM, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide" and a "composition comprising a given polypeptide" can refer to any composition containing the given polynucleotide or polypeptide. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotides encoding PMMM or fragments of PMMM may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (Accelrys, Burlington MA) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys

Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of PMMM or a polynucleotide encoding PMMM which can be identical in sequence to, but shorter in length than, the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from about 5 to about 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:59-116 can comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:59-116, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:59-116 can be employed in one or more embodiments of methods of the invention, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:59-116 from related polynucleotides. The precise length of a fragment of SEQ ID NO:59-116 and the region of SEQ ID NO:59-116 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-58 is encoded by a fragment of SEQ ID NO:59-116. A fragment of SEQ ID NO:1-58 can comprise a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-58. For example, a fragment of SEQ ID NO:1-58 can be used as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-58. The precise length of a fragment of SEQ ID NO:1-58 and the region of SEQ ID NO:1-58 to which the fragment corresponds can be determined based on the intended purpose for the fragment using one or more analytical methods described herein or otherwise known in the art.

A "full length" polynucleotide is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, alternatively, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of identical nucleotide matches between at least two polynucleotide sequences

aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using one or more computer algorithms or programs known in the art or described herein. For example, percent identity can be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989; CABIOS 5:151-153) and in Higgins, D.G. et al. (1992; CABIOS 8:189-191). For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms which can be used is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of identical residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide. The phrases "percent similarity" and "% similarity," as applied to polypeptide sequences, refer to the percentage of residue matches, including identical residue matches and conservative substitutions, between at least two polypeptide sequences aligned using a standardized algorithm. In contrast, conservative substitutions are not included in the calculation of percent identity between polypeptide sequences.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 μg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of

the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. and D.W. Russell (2001; Molecular Cloning: A Laboratory Manual, 3rd ed., vol. 1-3, Cold Spring Harbor Press, Cold Spring Harbor NY, ch. 9).

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acids by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid present in solution and another nucleic acid immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or polynucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of PMMM which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of PMMM which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, antibodies, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, antibody, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of PMMM. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of PMMM.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an PMMM may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of PMMM.

"Probe" refers to nucleic acids encoding PMMM, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acids. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100,

or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in, for example, Sambrook, J. and D.W. Russell (2001; Molecular Cloning: A Laboratory Manual, 3rd ed., vol. 1-3, Cold Spring Harbor Press, Cold Spring Harbor NY), Ausubel, F.M. et al. (1999; Short Protocols in Molecular Biology, 4th ed., John Wiley & Sons, New York NY), and Innis, M. et al. (1990; PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA). PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a nucleic acid that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook and Russell (*supra*). The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA molecule, is composed of the same linear sequence of nucleotides as the reference DNA molecule with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing PMMM, nucleic acids encoding PMMM, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide

comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably at least about 75% free, and most preferably at least about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" or "expression profile" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. In another embodiment, the nucleic acid can be introduced by infection with a recombinant viral vector, such as a lentiviral vector (Lois, C. et al. (2002) Science 295:868-872). The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods

known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook and Russell (*supra*).

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotides that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity or sequence similarity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity or sequence similarity over a certain defined length of one of the polypeptides.

THE INVENTION

Various embodiments of the invention include new human protein modification and maintenance molecules (PMMM), the polynucleotides encoding PMMM, and the use of these compositions for the diagnosis, treatment, or prevention of gastrointestinal, cardiovascular,

autoimmune/inflammatory, cell proliferative, developmental, epithelial, neurological, and reproductive disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide embodiments of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown. Column 6 shows the Incyte ID numbers of physical, full length clones corresponding to the polypeptide and polynucleotide sequences of the invention. The full length clones encode polypeptides which have at least 95% sequence identity to the polypeptide sequences shown in column 3.

Table 2 shows sequences with homology to polypeptide embodiments of the invention as identified by BLAST analysis against the GenBank protein (genpept) database and the PROTEOME database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (GenBank ID NO:) of the nearest GenBank homolog and the PROTEOME database identification numbers (PROTEOME ID NO:) of the nearest PROTEOME database homologs. Column 4 shows the probability scores for the matches between each polypeptide and its homolog(s). Column 5 shows the annotation of the GenBank and PROTEOME database homolog(s) along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Accelrys, Burlington MA). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are protein modification and maintenance

molecules. For example, SEQ ID NO:3 is 100% identical, from residue M1 to residue K696, to human ubiquitin-activating enzyme E1-like protein (GenBank ID g13623539) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:3 has homology to proteins that are ubiquitin-activating enzyme E1 type proteins, as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:3 also contains a ThiF and a ubiquitin-activating (UBA) protein repeat domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, BLAST and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:3 is a ubiquitin-activating enzyme. In an alternative example, SEQ ID NO:14 is 100% identical, from residue E162 to residue L345 and is 97% identical, from residue M1 to residue D165 to human neuroserpin (GenBank ID g1785654) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability scores are 5.5E-176, which indicate the probabilities of obtaining the observed polypeptide sequence alignments by chance. As determined by BLAST analysis using the PROTEOME database, SEQ ID NO:14 has homology to human and murine neuroserpin, members of the serpin family of serine protease inhibitors. Murine neuroserpin may attenuate extracellular proteolysis associated with neuronal migration, axogenesis, or synaptic connection formation during development. SEQ ID NO:14 also contains serpin and serine protease inhibitor domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, PROFILESCAN, and additional BLAST analyses provide further corroborative evidence that SEQ ID NO:14 is a neuroserpin. In an alternative example, SEQ ID NO:28 is 99% identical, from residue M1 to residue W124, and 93% identical, from residue C101 to residue K444, to human coagulation factor X (GenBank ID g182390, from M1 to F124 and from C140 to K488, respectively) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.6e-246, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:28 also has homology to proteins that are extracellular, have vitamin K-dependent serine protease activity, and are coagulation factor X proteins, as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:28 also contains trypsin, vitamin K-dependent carboxylation and EGF-like domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein families/domains. (See Table 3.) Data from BLIMPS, MOTIFS, BLASTand PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:28 is a coagulation

factor X protein. SEQ ID NO:1-2, SEQ ID NO:4-13, SEQ ID NO:15-27, and SEQ ID NO:29-58 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-58 are described in Table 7.

As shown in Table 4, the full length polynucleotide embodiments were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Column 1 lists the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5′) and stop (3′) positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide embodiments, and of fragments of the polynucleotides which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:59-116 or that distinguish between SEQ ID NO:59-116 and related polynucleotides.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotides. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as FL_$XXXXXX$_N_1_N_2_$YYYYY$_N_3_N_4 represents a "stitched" sequence in which *XXXXXX* is the identification number of the cluster of sequences to which the algorithm was applied, and *YYYYY* is the number of the prediction generated by the algorithm, and $N_{1,2,3...}$, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FL*XXXXXX*_g*AAAAA*_g*BBBBB*_1_*N* is a "stretched" sequence, with *XXXXXX* being the Incyte project identification number, g*AAAAA* being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, g*BBBBB* being the

GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and *N* referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, g*BBBBB*).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

Prefix	Type of analysis and/or examples of programs
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotides which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotides. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

Table 8 shows single nucleotide polymorphisms (SNPs) found in polynucleotide sequences of the invention, along with allele frequencies in different human populations. Columns 1 and 2 show the polynucleotide sequence identification number (SEQ ID NO:) and the corresponding Incyte project identification number (PID) for polynucleotides of the invention. Column 3 shows the Incyte identification number for the EST in which the SNP was detected (EST ID), and column 4 shows the identification number for the SNP (SNP ID). Column 5 shows the position within the EST sequence at which the SNP is located (EST SNP), and column 6 shows the position of the SNP within the full-length polynucleotide sequence (CB1 SNP). Column 7 shows the allele found in the EST sequence.

Columns 8 and 9 show the two alleles found at the SNP site. Column 10 shows the amino acid encoded by the codon including the SNP site, based upon the allele found in the EST. Columns 11-14 show the frequency of allele 1 in four different human populations. An entry of n/d (not detected) indicates that the frequency of allele 1 in the population was too low to be detected, while n/a (not available) indicates that the allele frequency was not determined for the population.

The invention also encompasses PMMM variants. Various embodiments of PMMM variants can have at least about 80%, at least about 90%, or at least about 95% amino acid sequence identity to the PMMM amino acid sequence, and can contain at least one functional or structural characteristic of PMMM.

Various embodiments also encompass polynucleotides which encode PMMM. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:59-116, which encodes PMMM. The polynucleotide sequences of SEQ ID NO:59-116, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses variants of a polynucleotide encoding PMMM. In particular, such a variant polynucleotide will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a polynucleotide encoding PMMM. A particular aspect of the invention encompasses a variant of a polynucleotide comprising a sequence selected from the group consisting of SEQ ID NO:59-116 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:59-116. Any one of the polynucleotide variants described above can encode a polypeptide which contains at least one functional or structural characteristic of PMMM.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide encoding PMMM. A splice variant may have portions which have significant sequence identity to a polynucleotide encoding PMMM, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to a polynucleotide encoding PMMM over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide encoding PMMM. For example, a polynucleotide comprising a sequence of SEQ ID NO:76, a polynucleotide

comprising a sequence of SEQ ID NO:77, a polynucleotide comprising a sequence of SEQ ID NO:78, a polynucleotide comprising a sequence of SEQ ID NO:89, a polynucleotide comprising a sequence of SEQ ID NO:90, and a polynucleotide comprising a sequence of SEQ ID NO:104 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:81 and a polynucleotide comprising a sequence of SEQ ID NO:82 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:86, a polynucleotide comprising a sequence of SEQ ID NO:87, and a polynucleotide comprising a sequence of SEQ ID NO:88 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:98 and a polynucleotide comprising a sequence of SEQ ID NO:115 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:101, a polynucleotide comprising a sequence of SEQ ID NO:103, and a polynucleotide comprising a sequence of SEQ ID NO:105 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:108, a polynucleotide comprising a sequence of SEQ ID NO:109, and a polynucleotide comprising a sequence of SEQ ID NO:110 are splice variants of each other; and a polynucleotide comprising a sequence of SEQ ID NO:112 and a polynucleotide comprising a sequence of SEQ ID NO:113 are splice variants of each other. Any one of the splice variants described above can encode a polypeptide which contains at least one functional or structural characteristic of PMMM.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding PMMM, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring PMMM, and all such variations are to be considered as being specifically disclosed.

Although polynucleotides which encode PMMM and its variants are generally capable of hybridizing to polynucleotides encoding naturally occurring PMMM under appropriately selected conditions of stringency, it may be advantageous to produce polynucleotides encoding PMMM or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding PMMM and its derivatives without altering the encoded amino acid sequences include the

production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of polynucleotides which encode PMMM and PMMM derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic polynucleotide may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a polynucleotide encoding PMMM or any fragment thereof.

Embodiments of the invention can also include polynucleotides that are capable of hybridizing to the claimed polynucleotides, and, in particular, to those having the sequences shown in SEQ ID NO:59-116 and fragments thereof, under various conditions of stringency (Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511). Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Biosciences, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Invitrogen, Carlsbad CA). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Amersham Biosciences), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art (Ausubel et al., *supra*, ch. 7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853).

The nucleic acids encoding PMMM may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). A

third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art (Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotides or fragments thereof which encode PMMM may be cloned in recombinant DNA molecules that direct expression of PMMM, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other polynucleotides which encode substantially the same or a functionally equivalent polypeptides may be produced and used to express PMMM.

The polynucleotides of the invention can be engineered using methods generally known in the art in order to alter PMMM-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA

shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of PMMM, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, polynucleotides encoding PMMM may be synthesized, in whole or in part, using one or more chemical methods well known in the art (Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232). Alternatively, PMMM itself or a fragment thereof may be synthesized using chemical methods known in the art. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques (Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J.Y. et al. (1995) Science 269:202-204). Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of PMMM, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography (Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421). The

composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing (Creighton, *supra*, pp. 28-53).

In order to express a biologically active PMMM, the polynucleotides encoding PMMM or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotides encoding PMMM. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of polynucleotides encoding PMMM. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where a polynucleotide sequence encoding PMMM and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

Methods which are well known to those skilled in the art may be used to construct expression vectors containing polynucleotides encoding PMMM and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination (Sambrook and Russell, *supra*, ch. 1-4, and 8; Ausubel et al., *supra*, ch. 1, 3, and 15).

A variety of expression vector/host systems may be utilized to contain and express polynucleotides encoding PMMM. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems (Sambrook and Russell, *supra*; Ausubel et al., *supra*; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc.

Natl. Acad. Sci. USA 81:3655-3659; Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355). Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of polynucleotides to the targeted organ, tissue, or cell population (Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5:350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Buller, R.M. et al. (1985) Nature 317:813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31:219-226; Verma, I.M. and N. Somia (1997) Nature 389:239-242). The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotides encoding PMMM. For example, routine cloning, subcloning, and propagation of polynucleotides encoding PMMM can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Invitrogen). Ligation of polynucleotides encoding PMMM into the vector's multiple cloning site disrupts the *lac*Z gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509). When large quantities of PMMM are needed, e.g. for the production of antibodies, vectors which direct high level expression of PMMM may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of PMMM. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign polynucleotide sequences into the host genome for stable propagation (Ausubel et al., *supra*; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184).

Plant systems may also be used for expression of PMMM. Transcription of polynucleotides encoding PMMM may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated

transfection (The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196).

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, polynucleotides encoding PMMM may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses PMMM in host cells (Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355).

For long term production of recombinant proteins in mammalian systems, stable expression of PMMM in cell lines is preferred. For example, polynucleotides encoding PMMM can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in tk^- and apr^- cells, respectively (Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823). Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14). Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites (Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051). Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β-

glucuronidase and its substrate β-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding PMMM is inserted within a marker gene sequence, transformed cells containing polynucleotides encoding PMMM can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding PMMM under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the polynucleotide encoding PMMM and that express PMMM may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of PMMM using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on PMMM is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art (Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding PMMM include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, polynucleotides encoding PMMM, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted

using a variety of commercially available kits, such as those provided by Amersham Biosciences, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with polynucleotides encoding PMMM may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode PMMM may be designed to contain signal sequences which direct secretion of PMMM through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted polynucleotides or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant polynucleotides encoding PMMM may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric PMMM protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of PMMM activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the PMMM encoding sequence and the heterologous protein sequence, so that PMMM may be cleaved away from the heterologous moiety following purification.

Methods for fusion protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In another embodiment, synthesis of radiolabeled PMMM may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

PMMM, fragments of PMMM, or variants of PMMM may be used to screen for compounds that specifically bind to PMMM. One or more test compounds may be screened for specific binding to PMMM. In various embodiments, 1, 2, 3, 4, 5, 10, 20, 50, 100, or 200 test compounds can be screened for specific binding to PMMM. Examples of test compounds can include antibodies, anticalins, oligonucleotides, proteins (e.g., ligands or receptors), or small molecules.

In related embodiments, variants of PMMM can be used to screen for binding of test compounds, such as antibodies, to PMMM, a variant of PMMM, or a combination of PMMM and/or one or more variants PMMM. In an embodiment, a variant of PMMM can be used to screen for compounds that bind to a variant of PMMM, but not to PMMM having the exact sequence of a sequence of SEQ ID NO:1-58. PMMM variants used to perform such screening can have a range of about 50% to about 99% sequence identity to PMMM, with various embodiments having 60%, 70%, 75%, 80%, 85%, 90%, and 95% sequence identity.

In an embodiment, a compound identified in a screen for specific binding to PMMM can be closely related to the natural ligand of PMMM, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner (Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2):Chapter 5). In another embodiment, the compound thus identified can be a natural ligand of a receptor PMMM (Howard, A.D. et al. (2001) Trends Pharmacol. Sci.22:132-140; Wise, A. et al. (2002) Drug Discovery Today 7:235-246).

In other embodiments, a compound identified in a screen for specific binding to PMMM can be closely related to the natural receptor to which PMMM binds, at least a fragment of the receptor, or a fragment of the receptor including all or a portion of the ligand binding site or binding pocket. For example, the compound may be a receptor for PMMM which is capable of propagating a signal, or a decoy receptor for PMMM which is not capable of propagating a signal (Ashkenazi, A. and V.M. Divit (1999) Curr. Opin. Cell Biol. 11:255-260; Mantovani, A. et al. (2001) Trends Immunol. 22:328-336). The compound can be rationally designed using known techniques. Examples of such techniques include those used to construct the compound etanercept (ENBREL; Amgen Inc.,

Thousand Oaks CA), which is efficacious for treating rheumatoid arthritis in humans. Etanercept is an engineered p75 tumor necrosis factor (TNF) receptor dimer linked to the Fc portion of human IgG_1 (Taylor, P.C. et al. (2001) Curr. Opin. Immunol. 13:611-616).

In one embodiment, two or more antibodies having similar or, alternatively, different specificities can be screened for specific binding to PMMM, fragments of PMMM, or variants of PMMM. The binding specificity of the antibodies thus screened can thereby be selected to identify particular fragments or variants of PMMM. In one embodiment, an antibody can be selected such that its binding specificity allows for preferential identification of specific fragments or variants of PMMM. In another embodiment, an antibody can be selected such that its binding specificity allows for preferential diagnosis of a specific disease or condition having increased, decreased, or otherwise abnormal production of PMMM.

In an embodiment, anticalins can be screened for specific binding to PMMM, fragments of PMMM, or variants of PMMM. Anticalins are ligand-binding proteins that have been constructed based on a lipocalin scaffold (Weiss, G.A. and H.B. Lowman (2000) Chem. Biol. 7:R177-R184; Skerra, A. (2001) J. Biotechnol. 74:257-275). The protein architecture of lipocalins can include a beta-barrel having eight antiparallel beta-strands, which supports four loops at its open end. These loops form the natural ligand-binding site of the lipocalins, a site which can be re-engineered *in vitro* by amino acid substitutions to impart novel binding specificities. The amino acid substitutions can be made using methods known in the art or described herein, and can include conservative substitutions (e.g., substitutions that do not alter binding specificity) or substitutions that modestly, moderately, or significantly alter binding specificity.

In one embodiment, screening for compounds which specifically bind to, stimulate, or inhibit PMMM involves producing appropriate cells which express PMMM, either as a secreted protein or on the cell membrane. Preferred cells can include cells from mammals, yeast, *Drosophila*, or *E. coli.* Cells expressing PMMM or cell membrane fractions which contain PMMM are then contacted with a test compound and binding, stimulation, or inhibition of activity of either PMMM or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with PMMM, either in solution or affixed to a solid support, and detecting the binding of PMMM to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical

libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

An assay can be used to assess the ability of a compound to bind to its natural ligand and/or to inhibit the binding of its natural ligand to its natural receptors. Examples of such assays include radio-labeling assays such as those described in U.S. Patent No. 5,914,236 and U.S. Patent No. 6,372,724. In a related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a receptor) to improve or alter its ability to bind to its natural ligands (Matthews, D.J. and J.A. Wells. (1994) Chem. Biol. 1:25-30). In another related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a ligand) to improve or alter its ability to bind to its natural receptors (Cunningham, B.C. and J.A. Wells (1991) Proc. Natl. Acad. Sci. USA 88:3407-3411; Lowman, H.B. et al. (1991) J. Biol. Chem. 266:10982-10988).

PMMM, fragments of PMMM, or variants of PMMM may be used to screen for compounds that modulate the activity of PMMM. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for PMMM activity, wherein PMMM is combined with at least one test compound, and the activity of PMMM in the presence of a test compound is compared with the activity of PMMM in the absence of the test compound. A change in the activity of PMMM in the presence of the test compound is indicative of a compound that modulates the activity of PMMM. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising PMMM under conditions suitable for PMMM activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of PMMM may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding PMMM or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease (see, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337). For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids

Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding PMMM may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding PMMM can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding PMMM is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress PMMM, e.g., by secreting PMMM in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of PMMM and protein modification and maintenance molecules. In addition, examples of tissues expressing PMMM can be found in Table 6 and can also be found in Example XI. Therefore, PMMM appears to play a role in gastrointestinal, cardiovascular, autoimmune/inflammatory, cell proliferative, developmental, epithelial, neurological, and reproductive disorders. In the treatment of disorders associated with increased PMMM expression or activity, it is desirable to decrease the expression or activity of PMMM. In the treatment of disorders associated with decreased PMMM expression or activity, it is desirable to increase the expression or activity of PMMM.

Therefore, in one embodiment, PMMM or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PMMM. Examples of such disorders include, but are not limited to, a gastrointestinal disorder, such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis,

gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha$_1$-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a cardiovascular disorder, such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation; an autoimmune/inflammatory disease, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, atherosclerotic plaque rupture, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, degradation of articular cartilage, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and

extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, colon, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a developmental disorder, such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, bone resorption, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, age-related macular degeneration, and sensorineural hearing loss; an epithelial disorder, such as dyshidrotic eczema, allergic contact dermatitis, keratosis pilaris, melasma, vitiligo, actinic keratosis, basal cell carcinoma, squamous cell carcinoma, seborrheic keratosis, folliculitis, herpes simplex, herpes zoster, varicella, candidiasis, dermatophytosis, scabies, insect bites, cherry angioma, keloid, dermatofibroma, acrochordons, urticaria, transient acantholytic dermatosis, xerosis, eczema, atopic dermatitis, contact dermatitis, hand eczema, nummular eczema, lichen simplex chronicus, asteatotic eczema, stasis dermatitis and stasis ulceration, seborrheic dermatitis, psoriasis, lichen planus, pityriasis rosea, impetigo, ecthyma, dermatophytosis, tinea versicolor, warts, acne vulgaris, acne rosacea, pemphigus vulgaris, pemphigus foliaceus, paraneoplastic pemphigus, bullous pemphigoid, herpes gestationis, dermatitis herpetiformis, linear IgA disease, epidermolysis bullosa acquisita, dermatomyositis, lupus erythematosus, scleroderma and morphea, erythroderma, alopecia, figurate skin lesions, telangiectasias, hypopigmentation, hyperpigmentation, vesicles/bullae, exanthems, cutaneous drug reactions, papulonodular skin lesions, chronic non-healing wounds, photosensitivity diseases, epidermolysis bullosa simplex, epidermolytic hyperkeratosis, epidermolytic and nonepidermolytic palmoplantar keratoderma, ichthyosis bullosa of Siemens, ichthyosis exfoliativa, keratosis palmaris et plantaris, keratosis palmoplantaris, palmoplantar keratoderma, keratosis punctata, Meesmann's corneal dystrophy, pachyonychia congenita, white sponge nevus, steatocystoma multiplex, epidermal nevi/epidermolytic hyperkeratosis type, monilethrix, trichothiodystrophy, chronic hepatitis/cryptogenic cirrhosis, and colorectal hyperplasia; a neurological disorder, such as epilepsy,

ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a reproductive disorder, such as infertility, including tubal disease, ovulatory defects, and endometriosis, a disorder of prolactin production, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia.

In another embodiment, a vector capable of expressing PMMM or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PMMM including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified PMMM in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PMMM including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of PMMM may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PMMM including, but not limited to, those listed above.

In a further embodiment, an antagonist of PMMM may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of PMMM. Examples of such disorders include, but are not limited to, those gastrointestinal, cardiovascular, autoimmune/inflammatory, cell proliferative, developmental, epithelial, neurological, and reproductive disorders described above. In one aspect, an antibody which specifically binds PMMM may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express PMMM.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding PMMM may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of PMMM including, but not limited to, those described above.

In other embodiments, any protein, agonist, antagonist, antibody, complementary sequence, or vector embodiments may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of PMMM may be produced using methods which are generally known in the art. In particular, purified PMMM may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind PMMM. Antibodies to PMMM may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. In an embodiment, neutralizing antibodies (i.e., those which inhibit dimer formation) can be used therapeutically. Single chain antibodies (e.g., from camels or llamas) may be potent enzyme inhibitors and may have application in the design of peptide mimetics, and in the development of immuno-adsorbents and biosensors (Muyldermans, S. (2001) J. Biotechnol. 74:277-302).

For the production of antibodies, various hosts including goats, rabbits, rats, mice, camels, dromedaries, llamas, humans, and others may be immunized by injection with PMMM or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are

not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to PMMM have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are substantially identical to a portion of the amino acid sequence of the natural protein. Short stretches of PMMM amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to PMMM may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce PMMM-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137).

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for PMMM may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and

easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 246:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between PMMM and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering PMMM epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for PMMM. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of PMMM-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple PMMM epitopes, represents the average affinity, or avidity, of the antibodies for PMMM. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular PMMM epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the PMMM-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of PMMM, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of PMMM-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available (Catty, *supra*; Coligan et al., *supra*).

In another embodiment of the invention, polynucleotides encoding PMMM, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA,

RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding PMMM. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding PMMM (Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press, Totawa NJ).

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein (Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102:469-475; Scanlon, K.J. et al. (1995) 9:1288-1296). Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors (Miller, A.D. (1990) Blood 76:271; Ausubel et al., *supra*; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63:323-347). Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art (Rossi, J.J. (1995) Br. Med. Bull. 51:217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87:1308-1315; Morris, M.C. et al. (1997) Nucleic Acids Res. 25:2730-2736).

In another embodiment of the invention, polynucleotides encoding PMMM may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in PMMM expression or regulation causes disease, the expression of PMMM from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in PMMM are treated by constructing mammalian expression vectors encoding PMMM and introducing these vectors by mechanical means into PMMM-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J.-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of PMMM include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). PMMM may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding PMMM from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to PMMM expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding PMMM under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are

commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In an embodiment, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding PMMM to cells which have one or more genetic abnormalities with respect to the expression of PMMM. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999; Annu. Rev. Nutr. 19:511-544) and Verma, I.M. and N. Somia (1997; Nature 18:389:239-242).

In another embodiment, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding PMMM to target cells which have one or more genetic abnormalities with respect to the expression of PMMM. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing PMMM to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which

consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999; J. Virol. 73:519-532) and Xu, H. et al. (1994; Dev. Biol. 163:152-161). The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another embodiment, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding PMMM to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for PMMM into the alphavirus genome in place of the capsid-coding region results in the production of a large number of PMMM-coding RNAs and the synthesis of high levels of PMMM in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of PMMM into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177). A

complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of RNA molecules encoding PMMM.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA molecules encoding PMMM. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

In other embodiments of the invention, the expression of one or more selected polynucleotides of the present invention can be altered, inhibited, decreased, or silenced using RNA interference (RNAi) or post-transcriptional gene silencing (PTGS) methods known in the art. RNAi is a post-transcriptional mode of gene silencing in which double-stranded RNA (dsRNA) introduced into a targeted cell specifically suppresses the expression of the homologous gene (i.e., the gene

bearing the sequence complementary to the dsRNA). This effectively knocks out or substantially reduces the expression of the targeted gene. PTGS can also be accomplished by use of DNA or DNA fragments as well. RNAi methods are described by Fire, A. et al. (1998; Nature 391:806-811) and Gura, T. (2000; Nature 404:804-808). PTGS can also be initiated by introduction of a complementary segment of DNA into the selected tissue using gene delivery and/or viral vector delivery methods described herein or known in the art.

RNAi can be induced in mammalian cells by the use of small interfering RNA also known as siRNA. SiRNA are shorter segments of dsRNA (typically about 21 to 23 nucleotides in length) that result *in vivo* from cleavage of introduced dsRNA by the action of an endogenous ribonuclease. SiRNA appear to be the mediators of the RNAi effect in mammals. The most effective siRNAs appear to be 21 nucleotide dsRNAs with 2 nucleotide 3' overhangs. The use of siRNA for inducing RNAi in mammalian cells is described by Elbashir, S.M. et al. (2001; Nature 411:494-498).

SiRNA can either be generated indirectly by introduction of dsRNA into the targeted cell, or directly by mammalian transfection methods and agents described herein or known in the art (such as liposome-mediated transfection, viral vector methods, or other polynucleotide delivery/introductory methods). Suitable SiRNAs can be selected by examining a transcript of the target polynucleotide (e.g., mRNA) for nucleotide sequences downstream from the AUG start codon and recording the occurrence of each nucleotide and the 3' adjacent 19 to 23 nucleotides as potential siRNA target sites, with sequences having a 21 nucleotide length being preferred. Regions to be avoided for target siRNA sites include the 5' and 3' untranslated regions (UTRs) and regions near the start codon (within 75 bases), as these may be richer in regulatory protein binding sites. UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP endonuclease complex. The selected target sites for siRNA can then be compared to the appropriate genome database (e.g., human, etc.) using BLAST or other sequence comparison algorithms known in the art. Target sequences with significant homology to other coding sequences can be eliminated from consideration. The selected SiRNAs can be produced by chemical synthesis methods known in the art or by *in vitro* transcription using commercially available methods and kits such as the SILENCER siRNA construction kit (Ambion, Austin TX).

In alternative embodiments, long-term gene silencing and/or RNAi effects can be induced in selected tissue using expression vectors that continuously express siRNA. This can be accomplished using expression vectors that are engineered to express hairpin RNAs (shRNAs) using methods known in the art (see, e.g., Brummelkamp, T.R. et al. (2002) Science 296:550-553; and Paddison, P.J. et al. (2002) Genes Dev. 16:948-958). In these and related embodiments, shRNAs can be delivered to target cells using expression vectors known in the art. An example of a suitable expression vector for

delivery of siRNA is the PSILENCER1.0-U6 (circular) plasmid (Ambion). Once delivered to the target tissue, shRNAs are processed *in vivo* into siRNA-like molecules capable of carrying out gene-specific silencing.

In various embodiments, the expression levels of genes targeted by RNAi or PTGS methods can be determined by assays for mRNA and/or protein analysis. Expression levels of the mRNA of a targeted gene, can be determined by northern analysis methods using, for example, the NORTHERNMAX-GLY kit (Ambion); by microarray methods; by PCR methods; by real time PCR methods; and by other RNA/polynucleotide assays known in the art or described herein. Expression levels of the protein encoded by the targeted gene can be determined by Western analysis using standard techniques known in the art.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding PMMM. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased PMMM expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding PMMM may be therapeutically useful, and in the treatment of disorders associated with decreased PMMM expression or activity, a compound which specifically promotes expression of the polynucleotide encoding PMMM may be therapeutically useful.

In various embodiments, one or more test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding PMMM is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding PMMM are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence

of the polynucleotide encoding PMMM. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art (Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466).

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of PMMM, antibodies to PMMM, and mimetics, agonists, antagonists, or inhibitors of PMMM.

In various embodiments, the compositions described herein, such as pharmaceutical compositions, may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery allows administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising PMMM or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, PMMM or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example PMMM or fragments thereof, antibodies of PMMM, and agonists, antagonists or inhibitors of PMMM, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity.

The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind PMMM may be used for the diagnosis of disorders characterized by expression of PMMM, or in assays to monitor patients being treated with PMMM or agonists, antagonists, or inhibitors of PMMM. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for PMMM include methods which utilize the antibody and a label to detect PMMM in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring PMMM, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of PMMM expression. Normal or standard values for PMMM expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to PMMM under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of PMMM expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, polynucleotides encoding PMMM may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotides, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of PMMM may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of PMMM, and to monitor regulation of PMMM levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotides, including genomic sequences, encoding PMMM or closely related molecules may be used to identify nucleic acid sequences which encode PMMM. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding PMMM, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the PMMM encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:59-116 or from genomic sequences including promoters, enhancers, and introns of the PMMM gene.

Means for producing specific hybridization probes for polynucleotides encoding PMMM include the cloning of polynucleotides encoding PMMM or PMMM derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotides encoding PMMM may be used for the diagnosis of disorders associated with expression of PMMM. Examples of such disorders include, but are not limited to, a gastrointestinal disorder, such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic

encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha$_1$-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a cardiovascular disorder, such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation; an autoimmune/inflammatory disease, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, atherosclerotic plaque rupture, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, degradation of articular cartilage, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, colon, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid,

penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a developmental disorder, such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, bone resorption, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, age-related macular degeneration, and sensorineural hearing loss; an epithelial disorder, such as dyshidrotic eczema, allergic contact dermatitis, keratosis pilaris, melasma, vitiligo, actinic keratosis, basal cell carcinoma, squamous cell carcinoma, seborrheic keratosis, folliculitis, herpes simplex, herpes zoster, varicella, candidiasis, dermatophytosis, scabies, insect bites, cherry angioma, keloid, dermatofibroma, acrochordons, urticaria, transient acantholytic dermatosis, xerosis, eczema, atopic dermatitis, contact dermatitis, hand eczema, nummular eczema, lichen simplex chronicus, asteatotic eczema, stasis dermatitis and stasis ulceration, seborrheic dermatitis, psoriasis, lichen planus, pityriasis rosea, impetigo, ecthyma, dermatophytosis, tinea versicolor, warts, acne vulgaris, acne rosacea, pemphigus vulgaris, pemphigus foliaceus, paraneoplastic pemphigus, bullous pemphigoid, herpes gestationis, dermatitis herpetiformis, linear IgA disease, epidermolysis bullosa acquisita, dermatomyositis, lupus erythematosus, scleroderma and morphea, erythroderma, alopecia, figurate skin lesions, telangiectasias, hypopigmentation, hyperpigmentation, vesicles/bullae, exanthems, cutaneous drug reactions, papulonodular skin lesions, chronic non-healing wounds, photosensitivity diseases, epidermolysis bullosa simplex, epidermolytic hyperkeratosis, epidermolytic and nonepidermolytic palmoplantar keratoderma, ichthyosis bullosa of Siemens, ichthyosis exfoliativa, keratosis palmaris et plantaris, keratosis palmoplantaris, palmoplantar keratoderma, keratosis punctata, Meesmann's corneal dystrophy, pachyonychia congenita, white sponge nevus, steatocystoma multiplex, epidermal nevi/epidermolytic hyperkeratosis type, monilethrix, trichothiodystrophy, chronic hepatitis/cryptogenic cirrhosis, and colorectal hyperplasia; a neurological disorder, such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal

familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a reproductive disorder, such as infertility, including tubal disease, ovulatory defects, and endometriosis, a disorder of prolactin production, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia. Polynucleotides encoding PMMM may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered PMMM expression. Such qualitative or quantitative methods are well known in the art.

In a particular embodiment, polynucleotides encoding PMMM may be used in assays that detect the presence of associated disorders, particularly those mentioned above. Polynucleotides complementary to sequences encoding PMMM may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of polynucleotides encoding PMMM in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of PMMM, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a

sequence, or a fragment thereof, encoding PMMM, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier, thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding PMMM may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding PMMM, or a fragment of a polynucleotide complementary to the polynucleotide encoding PMMM, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from polynucleotides encoding PMMM may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from polynucleotides encoding PMMM are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently

labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

SNPs may be used to study the genetic basis of human disease. For example, at least 16 common SNPs have been associated with non-insulin-dependent diabetes mellitus. SNPs are also useful for examining differences in disease outcomes in monogenic disorders, such as cystic fibrosis, sickle cell anemia, or chronic granulomatous disease. For example, variants in the mannose-binding lectin, MBL2, have been shown to be correlated with deleterious pulmonary outcomes in cystic fibrosis. SNPs also have utility in pharmacogenomics, the identification of genetic variants that influence a patient's response to a drug, such as life-threatening toxicity. For example, a variation in N-acetyl transferase is associated with a high incidence of peripheral neuropathy in response to the anti-tuberculosis drug isoniazid, while a variation in the core promoter of the ALOX5 gene results in diminished clinical response to treatment with an anti-asthma drug that targets the 5-lipoxygenase pathway. Analysis of the distribution of SNPs in different populations is useful for investigating genetic drift, mutation, recombination, and selection, as well as for tracing the origins of populations and their migrations (Taylor, J.G. et al. (2001) Trends Mol. Med. 7:507-512; Kwok, P.-Y. and Z. Gu (1999) Mol. Med. Today 5:538-543; Nowotny, P. et al. (2001) Curr. Opin. Neurobiol. 11:637-641).

Methods which may also be used to quantify the expression of PMMM include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves (Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotides described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function,

to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, PMMM, fragments of PMMM, or antibodies specific for PMMM may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time (Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484; hereby expressly incorporated by reference herein). Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression *in vivo*, as in the case of a tissue or biopsy sample, or *in vitro*, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with *in vitro* model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number

of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity (see, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm). Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In an embodiment, the toxicity of a test compound can be assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another embodiment relates to the use of the polypeptides disclosed herein to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed

by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of interest. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for PMMM to quantify the levels of PMMM expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art (Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/25116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662). Various types of microarrays are well known and thoroughly described in Schena, M., ed. (1999; DNA Microarrays: A Practical Approach, Oxford University Press, London).

In another embodiment of the invention, nucleic acid sequences encoding PMMM may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; Trask, B.J. (1991) Trends Genet. 7:149-154). Once mapped, the nucleic acid sequences may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP) (Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357).

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic map data (Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968). Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding PMMM on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to

11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation (Gatti, R.A. et al. (1988) Nature 336:577-580). The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, PMMM, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between PMMM and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest (Geysen, et al. (1984) PCT application WO84/03564). In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with PMMM, or fragments thereof, and washed. Bound PMMM is then detected by methods well known in the art. Purified PMMM can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding PMMM specifically compete with a test compound for binding PMMM. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PMMM.

In additional embodiments, the nucleotide sequences which encode PMMM may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/351,928, U.S. Ser. No. 60/359,903, and U.S. Ser. No. 60/366,837, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Invitrogen), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Invitrogen), using the recommended procedures or similar methods known in the art (Ausubel et al., *supra*, ch. 5). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Biosciences) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Invitrogen, Carlsbad CA), PCDNA2.1 plasmid (Invitrogen), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Invitrogen.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96

plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Biosciences or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Amersham Biosciences); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (Ausubel et al., *supra*, ch. 7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM; PROTEOME databases with sequences from *Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Candida albicans* (Incyte Genomics, Palo Alto CA); hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM (Haft, D.H. et al. (2001) Nucleic Acids Res. 29:41-43); and HMM-based protein domain databases such as SMART (Schultz, J. et al. (1998) Proc. Natl. Acad. Sci. USA 95:5857-5864; Letunic, I. et al. (2002)

Nucleic Acids Res. 30:242-244). (HMM is a probabilistic approach which analyzes consensus primary structures of gene families; see, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM; and HMM-based protein domain databases such as SMART. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (MiraiBio, Alameda CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:59-116. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative protein modification and maintenance molecules were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94; Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode protein modification and maintenance molecules, the encoded polypeptides were analyzed by querying against PFAM models for protein modification and maintenance molecules. Potential protein modification and maintenance molecules were also identified by homology to Incyte cDNA sequences that had been annotated as protein modification and maintenance molecules. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic

sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of PMMM Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:59-116 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:59-116 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

Association of PMMM polynucleotides with Lung Cancer

Heritable forms of lung carcinoma have not been reported and thus, identification of relevant disease-associated genes through conventional linkage analysis is not possible. However, several studies of sporadic nonsmall cell lung carcinoma (NSCLC) tumors have reported loss of heterozygosity (LOH) in regions of chromosome 11 suggesting the presence of one or more tumor suppressor genes (Bepler, G. and Garcia-Blanco, M.A. (1994) Proc. Natl. Acad. Sci. USA 91:5513-7; Iizuka, M. (1995) Genes, Chromosomes & Cancer 13:40-46; Rasio, D. (1995) Cancer Research 55:3988-91). In a study of 79 patients with lung cancer, Iizuka and coworkers found that 11q14-11q24.2 was deleted in many of the lung tumors studied. Mapping of this region with additional markers showed that the region of chromosome 11q bounded by markers D11S939 and D11S938 was commonly deleted (Iizuka, et al., *supra*). In another study it was shown that human A549 NSCLC cells transformed with a human-derived YAC clone containing a region of chromosome 11q within the region bounded by D11S939 and D11S938, exhibited little or no increase in cell number (versus control cells whose number increased 5-10-fold in the same 5 day period). Further studies of these hybrid cells showed a decrease in tumorigenicity and an increase in latency following injection into athymic, nude mice, as compared with mice injected with control A549 cells. These studies suggest the presence of a tumor suppressor gene within this region of chromosome 11q and support the association of LOH in this region with NSCLC.

Restriction fragment length polymorphism (RFLP) markers shown to be near regions of DNA known as sequence-tagged sites (STS), have been mapped to NT_Contigs generated by the Human Genome Project using ePCR (Schuler, G.D. (1997) Genome Research 7: 541-550, and (1998) Trends Biotechnol. 16(11):456-459). Contigs containing regions of DNA with known

disease-associated markers are therefore used to identify PMMM sequences that map to disease-associated regions of the genome.

Polynucleotides encoding PMMM were mapped to NT_Contigs. Contigs longer than 1Mb were broken into subcontigs of 1Mb length with overlapping sections of 100kb. A preliminary step used an algorithm, similar to MEGABLAST, to define the mRNA sequence /masked genomic DNA contig pairings. The cDNA/genomic pairings identified by the first algorithm were confirmed, and the PMMM polynucleotides mapped to DNA contigs, using SIM4 (Florea, L. et al. (1998) Genome Res. 8:967-74, version May 2000) which had been optimized for high throughput processing and strand assignment confidence). The SIM4 output of the mRNA sequence/genomic contig pairs was further processed to determine the correct location of the PMMM polynucleotides on the genomic contig, as well as their strand identity.

SEQ ID NO:66 was mapped to NT_Contig GBI:NT_009151_019.8; SEQ ID NO:76-77, SEQ ID NO:89-90, and SEQ ID NO:104 were mapped to NT_Contig GBI:NT_009151_022.8; and SEQ ID NO:96 was mapped to NT_Contig GBI:NT_009151_020.8 from Genbank, version 128, covering a 5.5 Mb region of the genome that also contains lung cancer-associated genetic markers D11S939 and D11S938. The maximum distance between SEQ ID NO:66, SEQ ID NO:76-77, SEQ ID NO:89-90, SEQ ID NO:104, and SEQ ID NO:96 and markers D11S939 and D11S938, therefore, is 5.5 Mb. Thus, SEQ ID NO:66, SEQ ID NO:76-77, SEQ ID NO:89-90, SEQ ID NO:104, and SEQ ID NO:96 are in proximity with genetic markers shown to consistently associate with lung cancer. Therefore, in various embodiments, SEQ ID NO:66, SEQ ID NO:76-77, SEQ ID NO:89-90, SEQ ID NO:104, and SEQ ID NO:96 can be used for one or more of the following: i) determination of LOH in persons with lung cancer in the lung cancer disease region at 11q12-24.2, ii) diagnostic assays for lung cancer, and iii) developing therapeutics and/or other treatments for lung cancer.

Association of PMMM polynucleotides with osteoarthritis

Markers that map to regions associated with particular diseases can be used to develop diagnostic and therapeutic tools. Disease association of a chromosome locus is expressed as the lod (logarithm of odds) score. The lod score is the logarithm to base 10 of the odds in favor of linkage. Linkage is defined as the tendency of two genes located on the same chromosome to be inherited together through meiosis (*Genetics in Medicine*, Fifth Edition, (1991) Thompson, M.W. et al. W.B. Saunders Co. Philadelphia). A logarithm of the odds ratio for linkage (lod) score of 2 indicates a probability of 1 in 100 that the marker was found solely by chance in affected individuals. In a study of 48 families affected by osteoarthritis (OA), Loughlin et al. (Rheumatology (2000) 39:377-381)

identified D2S202/D2S72 and D2S117 as two genetic markers with a multiple lod of 2.19 for linkage to OA of the hip.

Restriction fragment length polymorphism (RFLP) markers shown to be near regions of DNA known as sequence-tagged sites (STS), have been mapped to NT_Contigs generated by the Human Genome Project using ePCR (Schuler, G.D. (1997) Genome Research 7: 541-550, and (1998) Trends Biotechnol. 16(11):456-9). Contigs containing regions of DNA with known disease-associated markers are therefore used to identify PMMM sequences that map to disease-associated regions of the genome.

Polynucleotides encoding PMMM were mapped to NT_Contigs. Contigs longer than 1Mb were broken into subcontigs of 1Mb length with overlapping sections of 100kb. A preliminary step used an algorithm, similar to MEGABLAST, to define the mRNA sequence /masked genomic DNA contig pairings. The cDNA/genomic pairings identified by the first algorithm were confirmed, and the PMMM polynucleotides mapped to DNA contigs, using SIM4 (Florea, L. et al. (1998) Genome Res. 8:967-74, version May 2000) which had been optimized for high throughput processing and strand assignment confidence. The SIM4 output of the mRNA sequence/genomic contig pairs was further processed to determine the correct location of the PMMM polynucleotides on the genomic contig, as well as their strand identity.

SEQ ID NO:112-113 were mapped to NT_Contig GBI:NT_005229_002.8 from Genbank, version 128, covering a 6.45 Mb region of the genome that also contains OA-associated genetic markers D2S117 and D2S72. The maximum distance between SEQ ID NO:112-113 and markers D2S117 and D2S72, therefore, is 6.45 Mb. Thus, SEQ ID NO:112-113 are in proximity with genetic markers shown to consistently associate with OA. Therefore, in various embodiments, SEQ ID NO:112-113 can be used for one or more of the following: i) linkage analysis of persons and/or families to the OA disease region at 2q12-q22, ii) diagnostic assays for OA and interleukin expression abnormalities, and iii) developing therapeutics and/or other treatments for OA.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook and Russell, *supra*, ch. 7; Ausubel et al., *supra*, ch. 4).

Analogous computer techniques applying BLAST were used to search for identical or related molecules in databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer

search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score x Percent Identity}}{5 \text{ x minimum } \{\text{length(Seq. 1), length(Seq. 2)}\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotides encoding PMMM are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding PMMM. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of PMMM Encoding Polynucleotides

Full length polynucleotides are produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Biosciences), ELONGASE enzyme (Invitrogen), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Biosciences). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were

religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Biosciences), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Biosciences) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Biosciences) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotides are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Identification of Single Nucleotide Polymorphisms in PMMM Encoding Polynucleotides

Common DNA sequence variants known as single nucleotide polymorphisms (SNPs) were identified in SEQ ID NO:59-116 using the LIFESEQ database (Incyte Genomics). Sequences from the same gene were clustered together and assembled as described in Example III, allowing the identification of all sequence variants in the gene. An algorithm consisting of a series of filters was used to distinguish SNPs from other sequence variants. Preliminary filters removed the majority of basecall errors by requiring a minimum Phred quality score of 15, and removed sequence alignment errors and errors resulting from improper trimming of vector sequences, chimeras, and splice variants. An automated procedure of advanced chromosome analysis analysed the original chromatogram files in the vicinity of the putative SNP. Clone error filters used statistically generated algorithms to identify errors introduced during laboratory processing, such as those caused by reverse transcriptase, polymerase, or somatic mutation. Clustering error filters used statistically generated algorithms to identify errors resulting from clustering of close homologs or pseudogenes, or due to contamination by non-human sequences. A final set of filters removed duplicates and SNPs found in immunoglobulins or T-cell receptors.

Certain SNPs were selected for further characterization by mass spectrometry using the high throughput MASSARRAY system (Sequenom, Inc.) to analyze allele frequencies at the SNP sites in

four different human populations. The Caucasian population comprised 92 individuals (46 male, 46 female), including 83 from Utah, four French, three Venezualan, and two Amish individuals. The African population comprised 194 individuals (97 male, 97 female), all African Americans. The Hispanic population comprised 324 individuals (162 male, 162 female), all Mexican Hispanic. The Asian population comprised 126 individuals (64 male, 62 female) with a reported parental breakdown of 43% Chinese, 31% Japanese, 13% Korean, 5% Vietnamese, and 8% other Asian. Allele frequencies were first analyzed in the Caucasian population; in some cases those SNPs which showed no allelic variance in this population were not further tested in the other three populations.

X. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:59-116 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Biosciences), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Biosciences). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

XI. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing; see, e.g., Baldeschweiler et al., *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena, M., ed. (1999) DNA Microarrays: A Practical Approach, Oxford University Press, London). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface

of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements (Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31).

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Biosciences). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte Genomics). Specific control poly(A)$^+$ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (Clontech, Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Biosciences).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma-Aldrich, St. Louis MO) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1 μl of the array element DNA, at an average concentration of 100 ng/μl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μl of sample mixture consisting of 0.2 μg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is

focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte Genomics). Array elements that exhibit at least about a two-fold change in expression, a signal-to-background ratio of at least about 2.5, and an element spot size of at least about 40%, are considered to be differentially expressed.

Expression

For example, SEQ ID NO:62 showed differential expression associated with lung cancer, as determined by microarray analysis. Pair comparisons of lung tumor tissue and microscopically-normal tissue from the same donor were made. The expression of SEQ ID NO:62 was increased by at least two-fold in lung squamous cell carcinoma tissue from a 66 year-old male as compared to normal lung tissue from the same donor (Roy Castle International Centre for Lung Cancer Research, Liverpool, UK). Therefore, SEQ ID NO:13 is useful in diagnostic and disease staging assays for lung cancer and as a potential biological marker or therapeutic agents in the treatment of lung cancer.

In an alternative example, SEQ ID NO:67 showed differential expression in certain prostate carcinoma cell lines versus normal prostate epithelial cells as determined by microarray analysis. Three prostate carcinoma cell lines, DU 145, LNCaP, and PC-3 were included in the experiments. DU 145 was isolated from a metastatic site in the brain of a 69 year old male with widespread metastatic prostate carcinoma. DU 145 has no detectable sensitivity to hormones; forms colonies in semi-solid medium; is only weekly positive for acid phosphatase; and cells are negative for prostate specific antigen (PSA). LNCaP is a prostate carcinoma cell line isolated from a lymph node biopsy of a 50 year old male with metastatic prostate carcinoma. LNCaP expresses PSA, produces prostate acid phosphatase, and expresses androgen receptors. PC-3, a prostate adenocarcinoma cell line, was isolated from a metastatic site in the bone of a 62 year old male with grade IV prostate adenocarcinoma. The normal epithelial cell line, PrEC, is a primary prostate epithelial cell line isolated from a normal donor. In one experiment, the expression of cDNAs from the prostate carcinoma cell lines representing various stages of prostate tumor progression were compared with that of the normal prostate epithelial cells under the same culture conditions. The expression of cDNAs from the prostate carcinoma cell lines were compared to that of the normal prostate epithelial cells grown under the same conditions (in the absence of growth factors and hormones). The experiment showed that the expression of SEQ ID NO:67 was increased by at least two fold in DU145 prostate carcinoma lines relative to PrECs. Therefore, SEQ ID NO:67 is useful as a diagnostic marker or as a potential therapeutic target for certain prostate cancers.

In an alternative example, SEQ ID NO:69 also showed differential expression in certain prostate carcinoma cell lines versus normal prostate epithelial cells as determined by microarray analysis. The prostate carcinoma cell lines include CA-HPV-10, DU 145, LNCaP, and PC-3. CA-HPV-7 was derived from cells from a 63 year old male with prostate adenocarcinoma and was transformed by transfection with HPV18 DNA. DU 145 was isolated from a metastatic site in the brain of a 69 year old male with widespread metastatic prostate carcinoma. DU 145 has no detectable sensitivity to hormones; forms colonies in semi-solid medium; is only weakly positive for acid

phosphatase; and cells are negative for prostate specific antigen (PSA). LNCaP is a prostate carcinoma cell line isolated from a lymph node biopsy of a 50 year old male with metastatic prostate carcinoma. LNCaP expresses PSA, produces prostate acid phosphatase, and expresses androgen receptors. PC-3, a prostate adenocarcinoma cell line, was isolated from a metastatic site in the bone of a 62 year old male with grade IV prostate adenocarcinoma. The normal epithelial cell line, PZ-HPV-7 was derived from epithelial cells cultured from normal tissue from the peripheral zone of the prostate. The PZ-HPV-7 cells were transformed by transfection with HPV18. The microarray experiments showed that the expression of SEQ ID NO:69 was decreased by at least two fold in CA-HPV-10 prostate carcinoma line relative to cells from the normal prostate epithelial cell line, PZ-HPV-7. Therefore, SEQ ID NO:69 is useful as a diagnostic marker or as a potential therapeutic target for certain prostate cancers.

In an alternative example, the expression of SEQ ID NO:69 was increased by at least 2-fold in ovarian tumor tissue when matched with normal tissue from the same donor, a 79-year-old female donor with ovarian adenocarcinoma. Matched normal and tumorigenic ovarian tissue samples are provided by the Huntsman Cancer Institute, (Salt Lake City, UT). Therefore, SEQ ID NO:69 is useful in diagnostic assays and disease staging for ovarian cancer and as a potential biological marker and therapeutic agent in the treatment of ovarian cancer.

In an alternative example, SEQ ID NO:70 showed differential expression in certain breast carcinoma cell lines versus a non-malignant mammary epithelial cell line as determined by microarray analysis. The non-malignant mammary epithelial cell line, MCF10A, was isolated from a 36-year-old female with fibrocystic breast disease. The breast carcinoma cell lines include BT20, a breast carcinoma cell line derived in vitro from cells emigrating out of thin slices of a tumor mass isolated from a 74-year-old female; MCF7, a breast adenocarcinoma cell line derived from the pleural effusion of a 69-year-old female; MDA-mb-231, a metastatic breast tumor cell line derived from the pleural effusion of a 51-year-old female with metastatic breast carcinoma; SkBR3, a breast adenocarcinoma cell line isolated from a malignant pleural effusion of a 43-year-old female; and T47D, a breast carcinoma cell line derived from a pleural effusion from a 54-year-old female with an infiltrating ductal carcinoma of the breast. All cell cultures were propagated in a chemically-defined, serum-free medium, H14, according to the supplier's recommendations and grown to 70-80% confluence prior to RNA isolation. The microarray experiments showed that the expression of SEQ ID NO:70 was increased by at least two fold in BT20 breast carcinoma line relative to non-malignant mammary epithelial cells. Therefore, SEQ ID NO:70 is useful in diagnostic assays for and monitoring treatment of certain breast cancers.

In an alternative example, SEQ ID NO:94, SEQ ID NO:101, and SEQ ID NO:105, showed differential expression associated with breast cancer, as determined by microarray analysis. The gene expression profile of a nonmalignant mammary epithelial cell line was compared to the gene expression profiles of breast carcinoma lines at different stages of tumor progression. Cell lines compared included: a) BT-20, a breast carcinoma cell line derived *in vitro* from the cells emigrating out of thin slices of tumor mass isolated from a 74-year-old female, b) BT-474, a breast ductal carcinoma cell line that was isolated from a solid, invasive ductal carcinoma of the breast obtained from a 60-year-old woman, c) BT-483, a breast ductal carcinoma cell line that was isolated from a papillary invasive ductal tumor obtained from a 23-year-old normal, menstruating, parous female with a family history of breast cancer, d) Hs578T, a breast ductal carcinoma cell line isolated from a 74-year-old female with breast carcinoma, e) MCF7, a nonmalignant breast adenocarcinoma cell line isolated from the pleural effusion of a 69-year-old female, f) MCF-10A, a breast mammary gland (luminal ductal characteristics) cell line isolated. from a 36-year-old woman with fibrocystic breast disease, g) MDA-mb-435S, a spindle-shaped strain that evolved from the parent line (435) isolated by R. Cailleau from pleural effusion of a 31-year-old female with metastatic, ductal adenocarcinoma of the breast, h) Sk-BR-3, a breast adenocarcinoma cell line isolated from a malignant pleural effusion of a 43-year-old female, i) T-47D, a breast carcinoma cell line isolated from a pleural effusion obtained from a 54-year-old female with an infiltrating ductal carcinoma of the breast and j) HMEC, a primary breast epithelial cell line isolated from a normal donor. The expression of SEQ ID NO:94 was decreased by more than two-fold in carcinoma cell lines BT20, Sk-BR-3 and T-47D as compared to both HMEC cells and MCF10A cells. SEQ ID NO:101 expression was reduced by at least two-fold in BT20 and MCF7 cells as compared to HMEC cells. SEQ ID NO:105 expression was downregulated by at least two-fold in the carcinoma cell line BT20 as compared to the HMEC cell line. Therefore, SEQ ID NO:94, SEQ ID NO:101, and SEQ ID NO:105 are useful in monitoring treatment of, and diagnostic assays for breast cancer.

In an alternative example, SEQ ID NO:83, SEQ ID NO:101, SEQ ID NO:103, and SEQ ID NO:105 showed differential expression associated with lung cancer, as determined by microarray analysis. Expression was compared in matched samples of normal and lung tumor tissue from individual donors. Tissue samples were provided by the Roy Castle International Centre for Lung Cancer Research. The expression of SEQ ID NO:83 was decreased by at least two fold in squamous cell carcinoma tissue from a 73-year-old male donor and in lung adenocarcinoma tissue from a 71-year-old female donor, as compared to normal tissue from each of the respective donors. The expression of SEQ ID NO:101, SEQ ID NO:103, and SEQ ID NO:105 was upregulated by at least two-fold in three out of four lung tumor tissue samples as compared to normal tissue from each of the

four donors. Therefore, SEQ ID NO:83, SEQ ID NO:101, SEQ ID NO:103, and SEQ ID NO:105 are useful in monitoring treatment of, and diagnostic assays for lung cancer.

In an alternative example, SEQ ID NO:105 showed differential expression associated with colon cancer, as determined by microarray analysis. Matched normal and tumor samples from a 58-year-old female diagnosed with mucinous adenocarcinoma (Huntsman Cancer Institute, Salt Lake City, UT) were compared by competitive hybridization. SEQ ID NO:105 expression was upregulated by at least two-fold in the tumor tissue sample as compared to the normal tissue sample. Therefore, SEQ ID NO:105 is useful in monitoring treatment of, and diagnostic assays for colon cancer.

In an alternative example, SEQ ID NO:106 showed differential expression associated with prostate cancer, as determined by microarray analysis. Primary prostate epithelial cells were compared with prostate carcinomas representative of the different stages of tumor progression. Cell lines compared included: a)PrEC, a primary prostate epithelial cell line isolated from a normal donor, b)DU145, a prostate carcinoma cell line isolated from a metastatic site in the brain of 69-year old male with widespread metastatic prostate carcinoma, c)LNCaP, a prostate carcinoma cell line isolated from a lymph node biopsy of a 50-year-old male with metastatic prostate carcinoma, and d)PC-3, a prostate adenocarcinoma cell line isolated from a metastatic site in the bone of a 62- year-old male with grade IV prostate adenocarcinoma. Cells grown under restrictive conditions were compared to normal PrECs grown under restrictive conditions. Cells were grown in basal media in the absence of growth factors and hormones. The expression of SEQ ID NO:106 was increased by at least two-fold in DU145 cells as compared to PrEC cells. Therefore, SEQ ID NO:106 is useful in monitoring treatment of, and diagnostic assays for prostate cancer.

In an alternative example, SEQ ID NO:106 showed differential expression associated with Tangier disease (TD), as determined by microarray analysis. The expression of SEQ ID NO:106 was increased at least eight-fold in Tangier disease-derived fibroblasts compared to normal fibroblasts. In addition, both types of cells were cultured in the presence of cholesterol and compared with the same cell type cultured in the absence of cholesterol. Human fibroblasts were obtained from skin explants from both normal subjects and two patients with homozygous Tangier disease. Cell lines were immortalized by transfection with human papillomavirus 16 genes E6 and E7 and a neomycin resistance selectable marker. TD-derived cells are deficient in an assay of apoA-I mediated tritiated cholesterol efflux. Therefore, SEQ ID NO:106 is useful in monitoring treatment of, and diagnostic assays for Tangier disease.

In an alternative example, SEQ ID NO:106 showed differential expression associated with immune cells, as determined by microarray analysis. THP-1 is promonocyte cell line that was isolated from the peripheral blood of a 1-year-old male with acute monocytic leukemia. Upon stimulation with

PMA, THP-1 differentiates into a macrophage-like cell that displays many characteristics of peripheral human macrophages. THP-1 cells have been extensively used in the study of signaling in human monocytes and the identification of new factors produced by human monocytes. THP-1 cells were stimulated *in vitro* with soluble PMA, a broad activator of the protein kinase C-dependent pathways, and ionomycin, a calcium ionophore that permits the entry of calcium in the cell, for 0.5, 1, 2, 4, and 8 hours. These treated cells were compared to untreated THP-1 cells kept in culture in the absence of stimuli. At the 4 and 8 hour timepoints, the expression of SEQ ID NO:106 was increased by at least 3.5-fold in THP-1 cells treated with PMA and ionomycin as compared to untreated THP-1 cells. Jurkat cells, from an acute T cell leukemia cell line that grows actively in the absence of external stimuli and has been extensively used to study signaling in human T cells, were treated with combinations of graded doses of PMA and ionomycin and collected at a 1 hour time point. In T cells, the combination of PMA and ionomycin mimics the type of secondary signaling events elicited during optimal B cell activation. The treated cells were compared to untreated Jurkat cells kept in culture in the absence of stimuli. The expression of SEQ ID NO:106 was decreased by at least two-fold in Jurkat cells treated with 1 μM PMA and 1, 10, 50 or 200 μg/ml ionomycin. Therefore, SEQ ID NO:106 may be useful in monitoring of, and diagnostic assays for immune responses.

In an alternative example, SEQ ID NO:102 showed differential expression associated with obesity. Human primary subcutaneous preadipocytes were isolated from adipose tissue of a 36-year-old healthy female with a body mass index (BMI) of 27.7. The preadipocytes were cultured and induced to differentiate into adipocytes by culturing them in the differentiation medium containing the active components, PPAR-γ agonist and human insulin. Preadipocytes were treated with human insulin and PPAR-γ agonist for three days and subsequently were switched to medium containing insulin alone for a total duration of five days, 9 days or 12 days before the cells were collected for analysis. Differentiated adipocytes were compared to untreated preadipocytes maintained in culture in the absence of inducing agents. The expression of SEQ ID NO:102 was increased at least two-fold in treated human adipocytes from an obese donor when compared to non-treated adipocytes from the same donor. Thus, SEQ ID NO:102 is useful for the diagnosis, prognosis, or treatment of diabetes mellitus and other disorders such as obesity.

XII. Complementary Polynucleotides

Sequences complementary to the PMMM-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring PMMM. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of PMMM. To

inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the PMMM-encoding transcript.

XIII. Expression of PMMM

Expression and purification of PMMM is achieved using bacterial or virus-based expression systems. For expression of PMMM in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express PMMM upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of PMMM in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding PMMM by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus (Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945).

In most expression systems, PMMM is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Biosciences). Following purification, the GST moiety can be proteolytically cleaved from PMMM at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). Purified PMMM obtained by these methods can be used directly in the assays shown in Examples XVII, XVIII, XIX, and XX, where applicable.

XIV. Functional Assays

PMMM function is assessed by expressing the sequences encoding PMMM at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT plasmid (Invitrogen, Carlsbad CA) and PCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994; Flow Cytometry, Oxford, New York NY).

The influence of PMMM on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding PMMM and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding PMMM and other genes of interest can be analyzed by northern analysis or microarray techniques.

XV. Production of PMMM Specific Antibodies

PMMM substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize animals (e.g., rabbits, mice, etc.) and to produce antibodies using standard protocols.

Alternatively, the PMMM amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art (Ausubel et al., *supra*, ch. 11).

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity (Ausubel et al., *supra*). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-PMMM activity by, for example, binding the peptide or PMMM to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XVI. Purification of Naturally Occurring PMMM Using Specific Antibodies

Naturally occurring or recombinant PMMM is substantially purified by immunoaffinity chromatography using antibodies specific for PMMM. An immunoaffinity column is constructed by covalently coupling anti-PMMM antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Biosciences). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing PMMM are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PMMM (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/PMMM binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and PMMM is collected.

XVII. Identification of Molecules Which Interact with PMMM

PMMM, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent (Bolton, A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled PMMM, washed, and any wells with labeled PMMM complex are assayed. Data obtained using different concentrations of PMMM are used to calculate values for the number, affinity, and association of PMMM with the candidate molecules.

Alternatively, molecules interacting with PMMM are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989; Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

PMMM may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVIII. Demonstration of PMMM Activity

PMMM activity can be demonstrated using a generic immunoblotting strategy or through a variety of specific activity assays, some of which are outlined below. As a general approach, cell lines or tissues transformed with a vector containing PMMM coding sequences can be assayed for PMMM activity by immunoblotting. Transformed cells are denatured in SDS in the presence of b-mercaptoethanol, nucleic acids are removed by ethanol precipitation, and proteins are purified by acetone precipitation. Pellets are resuspended in 20 mM Tris buffer at pH 7.5 and incubated with Protein G-Sepharose pre-coated with an antibody specific for PMMM. After washing, the Sepharose beads are boiled in electrophoresis sample buffer, and the eluted proteins subjected to SDS-PAGE. The SDS-PAGE is transferred to a membrane for immunoblotting, and the PMMM activity is assessed by visualizing and quantifying bands on the blot using the antibody specific for PMMM as the primary antibody and ^{125}I-labeled IgG specific for the primary antibody as the secondary antibody.

PMMM kinase activity is measured by quantifying the phosphorylation of a protein substrate by PMMM in the presence of gamma-labeled ^{32}P-ATP. PMMM is incubated with the protein substrate, ^{32}P-ATP, and an appropriate kinase buffer. The ^{32}P incorporated into the substrate is separated from free ^{32}P-ATP by electrophoresis and the incorporated ^{32}P is counted using a radioisotope counter. The amount of incorporated ^{32}P is proportional to the activity of PMMM. A determination of the specific amino acid residue phosphorylated is made by phosphoamino acid analysis of the hydrolyzed protein.

PMMM phosphatase activity is measured by the hydrolysis of p-nitrophenyl phosphate (PNPP). PMMM is incubated together with PNPP in HEPES buffer, pH 7.5, in the presence of 0.1% β-mercaptoethanol at 37°C for 60 min. The reaction is stopped by the addition of 6 ml of 10 N NaOH and the increase in light absorbance at 410 nm resulting from the hydrolysis of PNPP is measured using a spectrophotometer. The increase in light absorbance is proportional to the activity of PMMM in the assay (Diamond, R.H. et al. (1994) Mol. Cell. Biol. 14:3752-3762).

In the alternative, PMMM phosphatase activity is determined by measuring the amount of phosphate removed from a phosphorylated protein substrate. Reactions are performed with 2 or 4 nM enzyme in a final volume of 30 μl containing 60 mM Tris, pH 7.6, 1 mM EDTA, 1 mM EGTA, 0.1% 2-mercaptoethanol and 10 μM substrate, ^{32}P-labeled on serine/threonine or tyrosine, as appropriate. Reactions are initiated with substrate and incubated at 30° C for 10-15 min. Reactions are quenched

with 450 μl of 4% (w/v) activated charcoal in 0.6 M HCl, 90 mM $Na_4P_2O_7$, and 2 mM NaH_2PO_4, then centrifuged at 12,000 × g for 5 min. Acid-soluble ^{32}Pi is quantified by liquid scintillation counting (Sinclair, C. et al. (1999) J. Biol. Chem. 274:23666-23672).

PMMM protease activity is measured by the hydrolysis of appropriate synthetic peptide substrates conjugated with various chromogenic molecules in which the degree of hydrolysis is quantified by spectrophotometric (or fluorometric) absorption of the released chromophore (Beynon, R.J. and J.S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford University Press, New York, NY, pp. 25-55). Peptide substrates are designed according to the category of protease activity as endopeptidase (serine, cysteine, aspartic proteases, or metalloproteases), aminopeptidase (leucine aminopeptidase), or carboxypeptidase (carboxypeptidases A and B, procollagen C-proteinase). Commonly used chromogens are 2-naphthylamine, 4-nitroaniline, and furylacrylic acid. Assays are performed at ambient temperature and contain an aliquot of the enzyme and the appropriate substrate in a suitable buffer. Reactions are carried out in an optical cuvette, and the increase/decrease in absorbance of the chromogen released during hydrolysis of the peptide substrate is measured. The change in absorbance is proportional to the enzyme activity in the assay.

In the alternative, an assay for PMMM protease activity takes advantage of fluorescence resonance energy transfer (FRET) that occurs when one donor and one acceptor fluorophore with an appropriate spectral overlap are in close proximity. A flexible peptide linker containing a cleavage site specific for PMMM is fused between a red-shifted variant (RSGFP4) and a blue variant (BFP5) of Green Fluorescent Protein. This fusion protein has spectral properties that suggest energy transfer is occurring from BFP5 to RSGFP4. When the fusion protein is incubated with PMMM, the substrate is cleaved, and the two fluorescent proteins dissociate. This is accompanied by a marked decrease in energy transfer which is quantified by comparing the emission spectra before and after the addition of PMMM (Mitra, R.D. et al (1996) Gene 173:13-17). This assay can also be performed in living cells. In this case the fluorescent substrate protein is expressed constitutively in cells and PMMM is introduced on an inducible vector so that FRET can be monitored in the presence and absence of PMMM (Sagot, I. et al (1999) FEBS Letters 447:53-57).

An assay for ubiquitin hydrolase activity measures the hydrolysis of a ubiquitin precursor. The assay is performed at ambient temperature and contains an aliquot of PMMM and the appropriate substrate in a suitable buffer. Chemically synthesized human ubiquitin-valine may be used as substrate. Cleavage of the C-terminal valine residue from the substrate is monitored by capillary electrophoresis (Franklin, K. et al. (1997) Anal. Biochem. 247:305-309).

PMMM protease inhibitor activity for alpha 2-HS-glycoprotein (AHSG) can be measured as a decrease in osteogenic activity in dexamethasone-treated rat bone marrow cell cultures (dex-RBMC).

Assays are carried out in 96-well culture plates containing minimal essential medium supplemented with 15% fetal bovine serum, ascorbic acid (50 mg/ml), antibiotics (100 mg/ml penicillin G, 50 mg/ml gentamicin, 0.3 mg/ml fungizone), 10 mM B-glycerophosphate, dexamethasone (10^{-8} M) and various concentrations of PMMM for 12-14 days. Mineralized tissue formation in the cultures is quantified by measuring the absorbance at 525 nm using a 96-well plate reader (Binkert, C. et al. (1999) J. Biol. Chem. 274:28514-28520).

PMMM protease inhibitor activity for inter-alpha-trypsin inhibitor (ITI) can be measured by a continuous spectrophotometric rate determination of trypsin activity. The assay is performed at ambient temperature in a quartz cuvette in pH 7.6 assay buffer containing 63 mM sodium phosphate, 0.23 mM N a-benzoyle-L-arginine ethyl ester, 0.06 mM hydrochloric acid, 100 units trypsin, and various concentrations of PMMM. Immediately after mixing by inversion, the increase in $A_{253\ nm}$ is recorded for approximately 5 minutes and the enzyme activity is calculated (Bergmeyer, H.U. et al. (1974) Meth. Enzym. Anal. 1:515-516).

PMMM isomerase activity such as peptidyl prolyl *cis/trans* isomerase activity can be assayed by an enzyme assay described by Rahfeld, J.U., et al. (1994; FEBS Lett. 352:180-184). The assay is performed at 10°C in 35 mM HEPES buffer, pH 7.8, containing chymotrypsin (0.5 mg/ml) and PMMM at a variety of concentrations. Under these assay conditions, the substrate, Suc-Ala-Xaa-Pro-Phe-4-NA, is in equilibrium with respect to the prolyl bond, with 80-95% in *trans* and 5-20% in *cis* conformation. An aliquot (2 ml) of the substrate dissolved in dimethyl sulfoxide (10 mg/ml) is added to the reaction mixture described above. Only the *cis* isomer of the substrate is a substrate for cleavage by chymotrypsin. Thus, as the substrate is isomerized by PMMM, the product is cleaved by chymotrypsin to produce 4-nitroanilide, which is detected by it's absorbance at 390 nm. 4-nitroanilide appears in a time-dependent and a PMMM concentration-dependent manner.

PMMM galactosyltransferase activity can be determined by measuring the transfer of radiolabeled galactose from UDP-galactose to a GlcNAc-terminated oligosaccharide chain (Kolbinger, F. et al. (1998) J. Biol. Chem. 273:58-65). The sample is incubated with 14 μl of assay stock solution (180 mM sodium cacodylate, pH 6.5, 1 mg/ml bovine serum albumin, 0.26 mM UDP-galactose, 2 μl of UDP-[^{3}H]galactose), 1 μl of $MnCl_2$ (500 mM), and 2.5 μl of GlcNAcβO-$(CH_2)_8$-CO_2Me (37 mg/ml in dimethyl sulfoxide) for 60 minutes at 37°C. The reaction is quenched by the addition of 1 ml of water and loaded on a C18 Sep-Pak cartridge (Waters), and the column is washed twice with 5 ml of water to remove unreacted UDP-[^{3}H]galactose. The [^{3}H]galactosylated GlcNAcβO-$(CH_2)_8$-CO_2Me remains bound to the column during the water washes and is eluted with 5 ml of methanol. Radioactivity in the eluted material is measured by liquid scintillation counting and is proportional to galactosyltransferase activity in the starting sample.

PMMM induction by heat or toxins may be demonstrated using primary cultures of human fibroblasts or human cell lines such as CCL-13, HEK293, or HEP G2 (ATCC). To heat induce PMMM expression, aliquots of cells are incubated at 42°C for 15, 30, or 60 minutes. Control aliquots are incubated at 37°C for the same time periods. To induce PMMM expression by toxins, aliquots of cells are treated with 100 μM arsenite or 20 mM azetidine-2-carboxylic acid for 0, 3, 6, or 12 hours. After exposure to heat, arsenite, or the amino acid analogue, samples of the treated cells are harvested and cell lysates prepared for analysis by western blot. Cells are lysed in lysis buffer containing 1% Nonidet P-40, 0.15 M NaCl, 50 mM Tris-HCl, 5 mM EDTA, 2 mM N-ethylmaleimide, 2 mM phenylmethylsulfonyl fluoride, 1 mg/ml leupeptin, and 1 mg/ml pepstatin. Twenty micrograms of the cell lysate is separated on an 8% SDS-PAGE gel and transferred to a membrane. After blocking with 5% nonfat dry milk/phosphate-buffered saline for 1 h, the membrane is incubated overnight at 4°C or at room temperature for 2-4 hours with an appropriate dilution of anti-PMMM serum in 2% nonfat dry milk/phosphate-buffered saline. The membrane is then washed and incubated with a 1:1000 dilution of horseradish peroxidase-conjugated goat anti-rabbit IgG in 2% dry milk/phosphate-buffered saline. After washing with 0.1% Tween 20 in phosphate-buffered saline, the PMMM protein is detected and compared to controls using chemiluminescence.

PMMM lysyl hydroxylase activity is determined by measuring the production of hydroxy[^{14}C]lysine from [^{14}C]lysine. Radiolabeled protocollagen is incubated with PMMM in buffer containing ascorbic acid, iron sulfate, dithiothreitol, bovine serum albumin, and catalase. Following a 30 minute incubation, the reaction is stopped by the addition of acetone, and centrifuged. The sedimented material is dried, and the hydroxy[^{14}C]lysine is converted to [^{14}C]formaldehyde by oxidation with periodate, and then extracted into toluene. The amount of ^{14}C extracted into toluene is quantified by scintillation counting, and is proportional to the activity of PMMM in the sample (Kivirikko, K., and R. Myllyla (1982) Methods Enzymol. 82:245-304).

XIX. Identification of PMMM Substrates

Phage display libraries can be used to identify optimal substrate sequences for PMMM. A random hexamer followed by a linker and a known antibody epitope is cloned as an N-terminal extension of gene III in a filamentous phage library. Gene III codes for a coat protein, and the epitope will be displayed on the surface of each phage particle. The library is incubated with PMMM under proteolytic conditions so that the epitope will be removed if the hexamer codes for a PMMM cleavage site. An antibody that recognizes the epitope is added along with immobilized protein A. Uncleaved phage, which still bear the epitope, are removed by centrifugation. Phage in the supernatant are then amplified and undergo several more rounds of screening. Individual phage clones are then isolated and sequenced. Reaction kinetics for these peptide substrates can be studied using an assay in

Example XVIII, and an optimal cleavage sequence can be derived (Ke, S.H. et al. (1997) J. Biol. Chem. 272:16603-16609).

To screen for *in vivo* PMMM substrates, this method can be expanded to screen a cDNA expression library displayed on the surface of phage particles (T7SELECT 10-3 Phage display vector, Novagen, Madison, WI) or yeast cells (pYD1 yeast display vector kit, Invitrogen, Carlsbad, CA). In this case, entire cDNAs are fused between Gene III and the appropriate epitope.

XX. Identification of PMMM Inhibitors

Compounds to be tested are arrayed in the wells of a multi-well plate in varying concentrations along with an appropriate buffer and substrate, as described in the assays in Example XVIII. PMMM activity is measured for each well and the ability of each compound to inhibit PMMM activity can be determined, as well as the dose-response kinetics. This assay could also be used to identify molecules which enhance PMMM activity.

In the alternative, phage display libraries can be used to screen for peptide PMMM inhibitors. Candidates are found among peptides which bind tightly to a protease. In this case, multi-well plate wells are coated with PMMM and incubated with a random peptide phage display library or a cyclic peptide library (Koivunen, E. et al. (1999) Nature Biotech 17:768-774). Unbound phage are washed away and selected phage amplified and rescreened for several more rounds. Candidates are tested for PMMM inhibitory activity using an assay described in Example XVIII.

Various modifications and variations of the described compositions, methods, and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. It will be appreciated that the invention provides novel and useful proteins, and their encoding polynucleotides, which can be used in the drug discovery process, as well as methods for using these compositions for the detection, diagnosis, and treatment of diseases and conditions. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Nor should the description of such embodiments be considered exhaustive or limit the invention to the precise forms disclosed. Furthermore, elements from one embodiment can be readily recombined with elements from one or more other embodiments. Such combinations can form a number of embodiments within the scope of the invention. It is intended that the scope of the invention be defined by the following claims and their equivalents.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
7313196	1	7313196CD1	59	7313196CB1	423513CA2
6465289	2	6465289CD1	60	6465289CB1	
7506357	3	7506357CD1	61	7506357CB1	
6878857	4	6878857CD1	62	6878857CB1	4995427CA2
7506021	5	7506021CD1	63	7506021CB1	
7503356	6	7503356CD1	64	7503356CB1	90033230CA2, 90033246CA2, 90124148CA2, 90124308CA2, 90124324CA2
7509052	7	7509052CD1	65	7509052CB1	1528449CA2, 2050062CA2, 3899930CA2, 4559732CA2, 4563755CA2, 4789452CA2, 56021587CA2, 6112175CA2, 8620878CA2, 90138573CA2
7503366	8	7503366CD1	66	7503366CB1	
7505933	9	7505933CD1	67	7505933CB1	
7507064	10	7507064CD1	68	7507064CB1	90112170CA2, 90112186CA2, 90112194CA2, 90112286CA2, 90112294CA2, 90115722CA2, 90115746CA2, 90115762CA2, 90115770CA2, 90115838CA2, 90115878CA2, 90115886CA2, 90116061CA2
1439986	11	1439986CD1	69	1439986CB1	
2008979	12	2008979CD1	70	2008979CB1	
90073157	13	90073157CD1	71	90073157CB1	90056580CA2, 90056588CA2, 90056596CA2, 90056664CA2, 90056672CA2
7506782	14	7506782CD1	72	7506782CB1	
7506941	15	7506941CD1	73	7506941CB1	
7507072	16	7507072CD1	74	7507072CB1	90123682CA2
7507083	17	7507083CD1	75	7507083CB1	95181896CA2
7509097	18	7509097CD1	76	7509097CB1	
7509118	19	7509118CD1	77	7509118CB1	
7509312	20	7509312CD1	78	7509312CB1	90134341CA2
90126902	21	90126902CD1	79	90126902CB1	
7509352	22	7509352CD1	80	7509352CB1	90003691CA2, 90124177CA2

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
7509341	23	7509341CD1	81	7509341CB1	90127811CA2
7509367	24	7509367CD1	82	7509367CB1	90127934CA2
7500455	25	7500455CD1	83	7500455CB1	5299346CA2, 90043663CA2, 90043779CA2, 95102091CA2
7510401	26	7510401CD1	84	7510401CB1	90040245CA2
7504702	27	7504702CD1	85	7504702CB1	1241883CA2, 90116231CA2, 90116378CA2, 90177975CA2, 90178051CA2, 90178095CA2, 90178111CA2, 90178119CA2
7509113	28	7509113CD1	86	7509113CB1	90134942CA2
7509140	29	7509140CD1	87	7509140CB1	90153388CA2
7509223	30	7509223CD1	88	7509223CB1	90135090CA2
7509272	31	7509272CD1	89	7509272CB1	90134126CA2
7509327	32	7509327CD1	90	7509327CB1	90134234CA2
7504677	33	7504677CD1	91	7504677CB1	95137671CA2
7504534	34	7504534CD1	92	7504534CB1	
7507771	35	7507771CD1	93	7507771CB1	
7504732	36	7504732CD1	94	7504732CB1	4921532CA2, 6824346CA2, 90005372CA2
950917	37	950917CD1	95	950917CB1	
7459720	38	7459720CD1	96	7459720CB1	
7503300	39	7503300CD1	97	7503300CB1	
7503334	40	7503334CD1	98	7503334CB1	90009616CA2
7503341	41	7503341CD1	99	7503341CB1	
7509936	42	7509936CD1	100	7509936CB1	
7509986	43	7509986CD1	101	7509986CB1	
7510010	44	7510010CD1	102	7510010CB1	
7510056	45	7510056CD1	103	7510056CB1	
7510398	46	7510398CD1	104	7510398CB1	90134242CA2
7510498	47	7510498CD1	105	7510498CB1	
7510044	48	7510044CD1	106	7510044CB1	
7504509	49	7504509CD1	107	7504509CB1	2678676CA2, 2679358CA2, 3608542CA2, 4331458CA2, 980724CA2

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
7506825	50	7506825CD1	108	7506825CB1	90115570CA2, 90115578CA2
7506828	51	7506828CD1	109	7506828CB1	
7506831	52	7506831CD1	110	7506831CB1	90115686CA2
7509968	53	7509968CD1	111	7509968CB1	95137452CA2
7510232	54	7510232CD1	112	7510232CB1	
7510233	55	7510233CD1	113	7510233CB1	
7510304	56	7510304CD1	114	7510304CB1	
7510461	57	7510461CD1	115	7510461CB1	3533809CA2
7510392	58	7510392CD1	116	7510392CB1	95139164CA2

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
1	7313196CD1	g183064	5.2E-167	[Homo sapiens] glia-derived nexin precursor Sommer, J. et al. Biochemistry 26:6407-6410 (1987)
		347042|SERPINE2	4.5E-168	[Homo sapiens][Inhibitor or repressor] Nexin, a member of the serpin family of serine protease inhibitors that is localized to neuromuscular junctions and to synapses in discrete regions of the brain, regulates alpha-thrombin Knauer, D. J. et al. J. Biol. Chem. 275:37340-6 (2000)
		588690|Spi4	1.9E-145	[Rattus norvegicus][Inhibitor or repressor; Protease (other than proteasomal)][Extracellular (excluding cell wall)] Glia-derived nexin, a neurite-promoting factor with serine protease inhibitory activity Smith-Swintosky, V. L. et al. J. Neurochem. 65:1415-8 (1995)
2	6465289CD1	g1429371	1.5E-132	[Drosophila melanogaster] ubiquitin-specific protease Henchoz, S. et al. Mol. Cell. Biol. 16:5717-5725 (1996)
		249849|T05H10.1	1.2E-100	[Caenorhabditis elegans][Hydrolase; Protease (other than proteasomal)] Putative ubiquitin-specific protease (ubiquitin C-terminal hydrolase), possible ortholog of D. melanogaster UBP64E ubiquitin-specific protease 64E
		658508|ubpD	3.9E-61	[Schizosaccharomyces pombe] Putative ubiquitin carboxyl-terminal hydrolase
3	7506357CD1	g13623539	0.0	[Homo sapiens] ubiquitin-activating enzyme E1-like
		338764|UBE1L	0.0	[Homo sapiens][Activator; Protein conjugation factor] Ubiquitin-activating enzyme E1-like, catalyzes the activation step in the conjugation of ISG15, a ubiquitin-like protein induced by interferon; corresponding gene is located in a chromosomal region associated with loss of heterozygosity in lung cancer McLaughlin, P. M. et al. Int. J. Cancer 85:871-6 (2000)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		581777|Ube1x	7.4E-225	[Mus musculus][Activator; Protein conjugation factor] Ubiquitin-activating enzyme E1, a protein that activates ubiquitin to mark cellular proteins for degradation, may play a role in DNA repair Imai, N. et al. Gene 118:279-82 (1992)
4	6878857CD1	g521218	9.1E-120	[Homo sapiens] trypsinogen Emi, M. et al. Gene 41:305-310 (1986)
		337264|PRSS2	7.7E-121	[Homo sapiens][Hydrolase; Protease (other than proteasomal)] Trypsinogen-2 (pancreatic trypsin II, anionic trypsinogen), a serine protease produced mainly in the pancreas, presence in serum or urine is used as a marker for pancreatitis Hedstrom, J. et al. Am. J. Gastroenterol. 96:424-30 (2001)
		337262|PRSS1	2.2E-108	[Homo sapiens][Hydrolase; Protease (other than proteasomal)][Extracellular (excluding cell wall)] Pancreatic trypsinogen 1, inactive precursor of the serine protease trypsin that cleaves peptide bonds at the carboxyl side of Lysine or Arginine and is involved in digestion; mutations in the corresponding gene are associated with hereditary pancreatitis Sahin-Toth, M. and Toth, M. Biochem. Biophys. Res. Commun. 278:286-9 (2000)
5	7506021CD1	g14582773	2.0E-226	[Homo sapiens] sumo/sentrin-specific protease
		569442|SENP1	1.4E-28	[Homo sapiens][Hydrolase; Protease (other than proteasomal)][Nuclear] Sentrin/SUMO-specific protease, an endopeptidase localized to the nucleus which selectively removes sentrin from protein-sentrin conjugates such as those formed by the tumor suppressor protein PML Gong, L. et al. J. Biol. Chem. 275:3355-9 (2000)
6	7503356CD1	g9802310	2.2E-26	[Homo sapiens] calpain large polypeptide L2 (Ye, Z. and Connor, J. R. (2000) Biochem. Biophys. Res. Commun. 275:223-227.)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		334452|CAPN2	2.4E-27	[Homo sapiens][Hydrolase; Protease (other than proteasomal)] Calpain 2, large subunit of the cysteine-type protease m-calpain which may regulate the cell cycle, apoptosis, and cellular differentiation, upregulated in muscle from progressive muscular dystrophy and amyotrophic lateral sclerosis patients (Ueyama, H. et al. (1998) J. Neurol. Sci. 155:, 163-169; Hata, S. et al. (2001) FEBS Lett. 501:111-114.)
		590923|Capn2	6.5E-27	[Rattus norvegicus][Hydrolase; Protease (other than proteasomal)][Cytoplasmic] Calpain 2, large subunit of the cysteine-type protease m-calpain which regulates apoptosis and signal transduction; human CAPN2 is elevated in muscle from progressive muscular dystrophy and amyotrophic lateral sclerosis patients (Posmantur, R. et al. (1997) Neuroscience 77:875-888; Hosfield, C. M. et al. (2001) J. Biol. Chem. 276:7404-7407.)
8	7503366CD1	g14279329	0.0	[Homo sapiens] ubiquitin specific protease
		432832|USP25	3.0E-152	[Homo sapiens][Hydrolase; Protease (other than proteasomal)] Ubiquitin specific protease 25, a C-terminal ubiquitin hydrolase; loss of heterozygosity is seen in non small cell lung carcinomas, candidate for involvement in chromosome 21 Trisomy (Down syndrome) and its associated defective spermiogenesis (Valero, R. et al. (1999) Genomics 62:395-405; Groet, J. et al. (2000) Genes Chromosomes Cancer 27:153-161.)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		438259|Usp25	3.0E-150	[Mus musculus][Hydrolase;Protease (other than proteasomal)] Ubiquitin specific protease 25, putative C-terminal ubiquitin hydrolase, may be involved in the development and differentiation of highly proliferative tissues; human USP25 shows loss of heterozygosity in non small cell lung carcinomas (Valero, R. et al. (1999) supra.)
9	7505933CD1	g1256401	9.5E-183	[Homo sapiens] 26S protease subunit S5a (Ferrell, K. et al. (1996) FEBS Lett. 381:143-148.)
		343294|PSMD4	8.0E-184	[Homo sapiens][Proteasome subunit][Cytoplasmic] Proteasome 26S subunit non ATPase 4, an antisecretory factor that is a subunit of the 26S proteasome and may bind to multiubiquitinated proteins; inhibits intestinal fluid secretion induced by cholera toxin (Tanahashi, N. et al. (2000) J. Biol. Chem. 275:14336-14345; Lange, S. et al. (1999) Cell Tissue Res. 296:607-617.)
		711344|Psmd4	5.1E-180	[Rattus norvegicus][Proteasome subunit][Cytoplasmic] Proteasome 26S subunit non ATPase 4, an antisecretory factor that is a subunit of the 26S proteasome and may bind to multiubiquitinated proteins; inhibits cholera toxin-induced secretion of intestinal fluid (Tateishi, K. et al. (1999) Biochem. Cell Biol. 77:223-228; Noda, C. et al. (2000) Biochem. Biophys. Res. Commun. 277:348-354.)
10	7507064CD1	g7106000	6.5E-45	[Rattus norvegicus] prochymosin (Kageyama, T. et al. (2000) Biochem. Biophys. Res. Commun. 267:806-812)
		610015|LOC56825	5.4E-46	[Rattus norvegicus][Hydrolase; Protease (other than proteasomal)][Extracellular (excluding cell wall)] Prochymosin, a major neonatal pepsinogen, expressed only at neonatal-infant stages (Kageyama, T. et al. (2000) supra.)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		724632|5pep_	3.9E-32	[Protein Data Bank] Pepsin (E.C. 3.4.23.1) (Cooper, J. B. et al. (1990) J. Mol. Biol. 214:199-222.)
11	1439986CD1	g3603463	9.1E-78	[Mus musculus] heat shock protein hsp40-3 (Chen, M. S. et al. (1999) Gene 238:333-341.)
		428796|DNAJB5	8.7E-80	[Homo sapiens][Chaperones;Heat shock protein] DnaJ (Hsp40) homolog subfamily B member 5 (heat shock cognate 40), member of the Hsp40 heat shock protein family, putative chaperone predicted to be involved in the stress response, may be a specific functional partner for HSC70 (HSPA8) (Chen, M. S. et al. (1999) Gene 238:333-341.)
		609024|Dnajb5	7.7E-79	[Mus musculus][Chaperones;Heat shock protein] DnaJ (Hsp40) homolog subfamily B member 5 (heat shock cognate 40), member of the Hsp40 heat shock protein family, putative chaperone predicted to be involved in the stress response, may be a specific functional partner for Hsc70 (Hspa8) (Chen, M. S. et al. (1999) supra.)
12	2008979CD1	g12324903	6.0E-53	[Arabidopsis thaliana] putative heat shock protein; 32627-30541
		249587|T03F6.2	1.4E-74	[Caenorhabditis elegans][Chaperones] Protein containing a DnaJ domain, has strong similarity to S. cerevisiae Ynl227p (Jiang, M. et al. (2001) Proc. Natl. Acad. Sci. U S A 98:218-223;).
		8419|YNL227C	1.1E-48	[Saccharomyces cerevisiae][Chaperones] Member of the DnaJ family of protein chaperones, contains two C2H2-type zinc finger domains (Bohm, S. et al. (1997) Nucleic Acids Res 25:2464-2469; Ito, T. et al. (2001) Proc. Natl. Acad. Sci. U S A 98:4569-4574.)
13	90073157CD1	g180503	2.8E-55	[Homo sapiens] di-N-acetylchitobiase (Fisher, K. J. and Aronson, N. N. Jr. (1992) J. Biol. Chem. 267:19607-19616.)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		340298|CTBS	2.4E-56	[Homo sapiens][Hydrolase][Lysosome/vacuole;Cytoplasmic] Chitobiase (di-N-acetylchitobiase), a lysosomal glycosidase involved in the degradation of asparagine-linked glycoproteins (Fisher, K. J., and Aronson, N. N. Jr. (1992) Mol. Cell Biol. 12:1585-1591.)
		704880|Ctbs	9.7E-35	[Rattus norvegicus][Hydrolase][Lysosome/vacuole;Cytoplasmic] Chitobiase (di-N-acetylchitobiase), a lysosomal glycosidase involved in the degradation of asparagine-linked glycoproteins (Kuranda, M. J., and Aronson, N. N. Jr. (1986) J. Biol. Chem. 261:5803-5809.)
14	7506782CD1	g1785654	5.5E-176	[Homo sapiens] neuroserpin (Schrimpf, S. P. et al. (1997) Genomics 40:55-62.)
		341770|SERPINI1	4.7E-177	[Homo sapiens][Inhibitor or repressor] Neuroserpin, a member of the serpin family of serine protease inhibitors; chicken homolog is secreted by the axons of central and peripheral nervous system neurons (Davis, R. L. et al. (1999) Nature 401:376-379; Hastings, G. A. et al. (1997) J. Biol. Chem. 272:33062-33067.)
		581457|Serpini1	4.5E-155	[Mus musculus][Inhibitor or repressor] Neuroserpin, an inhibitor of the serine protease tissue plasminogen activator, may attenuate extracellular proteolysis associated with neuronal migration, axogenesis, or synaptic connection formation during development (Krueger, S. R. et al. (1997) J. Neurosci. 17:8984-8996; Hastings, G. A. et al. (1997) supra.)
15	7506941CD1	g11095188	7.2E-220	[Homo sapiens] dipeptidyl peptidase 8 (Abbott, C. A. et al. (2000) Eur. J. Biochem. 267: 6140-6150.)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		728646|K02F2.1	2.1E-44	[Caenorhabditis elegans][Protease (other than proteasomal)][Golgi] Putative peptidase, has strong similarity to S. cerevisiae STE13 gene product [aminopeptidase A] (Bateman, A. et al. (1999) Nucleic Acids Res. 27:260-262; Jiang, M. et al. (2001) Proc. Natl. Acad. Sci. U S A 98:218-223.)
16	7507072CD1	g17226093	3.0E-90	[Homo sapiens] (AF384667) carboxypeptidase A5
		590975|Cpa1	2.1E-41	[Rattus norvegicus][Hydrolase;Protease (other than proteasomal)] Pancreatic procarboxypeptidase A1, a secreted digestive metalloprotease that hydrolyzes the C-terminal peptide bond of polypeptides with specificity for C-terminal residues with aromatic and branched aliphatic side chains (Gardell, S. J. et al. (1988) J. Biol. Chem. 263:17828-17836; Clauser, E. et al. (1988) J. Biol. Chem. 263:17837-17845.)
		342304|CPA1	5.5E-41	[Homo sapiens][Hydrolase;Protease (other than proteasomal)] Pancreatic procarboxypeptidase A1, a secreted monomeric digestive metalloprotease that hydrolyzes the C-terminal peptide bond of polypeptides with specificity for C-terminal residues with aromatic and branched aliphatic side chains (Catasus, L. et al. (1992) Biochem. J. 287:299-303; Garcia-Saez, I. et al. (1997) EMBO J. 16:6906-6913.)
17	7507083CD1	g16588795	3.2E-15	[Rattus norvegicus] (AF317633) capsase-12
		584399|Casp12	4.4E-16	[Mus musculus][Hydrolase;Protease (other than proteasomal)][Endoplasmic reticulum;Cytoplasmic] Caspase 12, a member of the caspase family of cysteine proteases, a caspase that induces apoptosis in response to endoplasmic reticulum-specific stress, may contribute to amyloid beta neurotoxicity in cortical neurons (Nakagawa, T., and Yuan, J. (2000) J. Cell Biol. 150:887-894; Nakagawa, T. et al. (2000) Nature 403:98-103.)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		339702|CASP13	3.1E-11	[Homo sapiens][Hydrolase;Protease (other than proteasomal)] Caspase 13, a cysteine (thiol) protease, member of the ICE (CASP1) subfamily of caspases, activated by caspase 8 (CASP8), induces apoptosis when overexpressed and has a role in tumor necrosis factor-induced apoptosis (Humke, E. W. et al. (1998) J. Biol. Chem. 273:15702-15707; Lin, X. Y. et al. (2000) J. Biol. Chem. 275:39920-39926.)
18	7509097CD1	g14348014	7.4E-126	[Homo sapiens] DNA encoding a transmembrane serine protease (Endotheliase 2-S) protein
		742472|MSP	2.8E-108	[Homo sapiens] Mosaic serine protease, contains a serine protease domain and regulatory modules having a protein kinase substrate and a low-density lipoprotein receptor (Kim, D. R. et al. (2001) Biochim. Biophys. Acta 1518:204-209.)
19	7509118CD1	g14348014	2.4E-207	[Homo sapiens] DNA encoding a transmembrane serine protease (Endotheliase 2-S) protein
		742472|MSP	1.4E-188	[Homo sapiens] Mosaic serine protease, contains a serine protease domain and regulatory modules having a protein kinase substrate and a low-density lipoprotein receptor (Kim, D. R. et al. (2001) supra.)
		588075|Tmprss2	1.6E-34	[Mus musculus][Hydrolase;Protease (other than proteasomal)][Plasma membrane] Epitheliasin, a serine proteinase that contains transmembrane, LDLRA (LDL receptor class A) and SRCR (scavenger receptor cysteine-rich) domains; human TMPRSS2 is highly expressed in androgen-dependent prostate cancer (Jacquinet, E. et al. (2000) FEBS Lett. 468:93-100.)
20	7509312CD1	g14348014	7.8E-202	[Homo sapiens] DNA encoding a transmembrane serine protease (Endotheliase 2-S) protein

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		742472|MSP	2.0E-162	[Homo sapiens] Mosaic serine protease, contains a serine protease domain and regulatory modules having a protein kinase substrate and a low-density lipoprotein receptor (Kim, D. R. et al. (2001) supra.)
		588075|Tmprss2	2.7E-39	[Mus musculus][Hydrolase;Protease (other than proteasomal)][Unspecified membrane;Plasma membrane] Epitheliasin, a serine proteinase that contains transmembrane, LDLRA (LDL receptor class A) and SRCR (scavenger receptor cysteine-rich) domains; human TMPRSS2 is highly expressed in androgen-dependent prostate cancer (Jacquinet, E. et al. (2000) supra.)
21	90126902CD1	g11935122	2.8E-52	[Mus musculus] papilin (Kramerova, I. A. et al. (2000) Development 127: 5475-5485.)
		568954|ADAMTS6	6.0E-52	[Homo sapiens][Hydrolase;Protease (other than proteasomal)] Disintegrin like metalloprotease with thrombospondin type 1 motif 6, a member of the ADAMTS family of disintegrin metalloproteases with thrombospondin motifs, has a cysteine-rich and a reprolysin-type catalytic domain (Hurskainen, T. L. et al. (1999) J. Biol. Chem. 274:25555-25563)
		341526|ADAMTS4	3.0E-43	[Homo sapiens][Hydrolase;Protease (other than proteasomal)][Unspecified membrane] Aggrecanase-1, a zinc metalloproteinase member of a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS) family, cleaves and proteolyzes cartilage aggrecan and may play a role in aggrecan degradation observed in arthritic diseases (Tang, B. L. (2001) Int. J. Biochem. Cell Biol. 33:33-44; Tortorella, M. D. et al. (1999) Science 284:1664-1666.)
22	7509352CD1	g14669471	5.8E-146	[Homo sapiens] lysyl oxidase-related protein C Ito, H. et al. J. Biol. Chem. 276, 24023-24029 (2001)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		758804|LOXL4	4.9E-147	[Homo sapiens] Member of the lysyl oxidase (LOX) family, which catalyze the oxidative deamination of collagen peptidyl lysine residues, has strong similarity to a region of lysyl oxidase-like 4 (mouse Loxl4), which plays a role in extracellular matrix formation Ito, et al. *(supra)*
		753883|Loxl4	3.5E-127	[Mus musculus] Lysyl oxidase-like 4, catalyzes the oxidative deamination of peptidyl lysine residues and protein crosslinking during collagenous extracellular matrix formation, involved in chondrocyte differentiation; activity is inhibited by beta-aminopropionitrile
23	7509341CD1	g5921501	1.7E-82	[Mus musculus] distal intestinal serine protease Shaw-Smith, C. J. et al. Biochim. Biophys. Acta 1490, 131-136 (2000)
		438199|Disp	1.4E-83	[Mus musculus] [Hydrolase; Protease (other than proteasomal)] Distal intestinal serine protease, member of the serine protease clan SA/family S1, expressed predominantly in distal gut Shaw-Smith, C. J. et al. (supra)
		709655|MPN	5.5E-48	[Homo sapiens] Member of the trypsin family of serine proteases, has moderate similarity to prostasin (human PRSS8), which is a transmembrane serine protease that is proteolytically released from the membrane upon secretion
24	7509367CD1	g5921501	1.5E-26	[Mus musculus] distal intestinal serine protease Shaw-Smith, C. J. et al. (supra)
		438199|Disp	1.2E-27	[Mus musculus] [Hydrolase; Protease (other than proteasomal)] Distal intestinal serine protease, member of the serine protease clan SA/family S1, expressed predominantly in distal gut Shaw-Smith, C. J. et al. (supra)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		709655|MPN	1.6E-16	[Homo sapiens] Member of the trypsin family of serine proteases, has moderate similarity to prostasin (human PRSS8), which is a transmembrane serine protease that is proteolytically released from the membrane upon secretion
26	7510401CD1	570808|LOC51035	5.6E-87	[Homo sapiens] Protein containing a UBX domain, which are found in ubiquitin regulatory proteins, contains a UBA/THIF-type NAD/FAD binding fold
27	7504702CD1	g4235425	3.3E-36	[Homo sapiens] napsin 1 precursor
		340906|NAP1	2.8E-37	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] Pronapsin A, aspartic-type endopeptidase that may play a role in lung surfactant precursor proteolytic processing and as a marker for primary lung adenocarcinoma; lack of NAP1 in the urine is associated with kidney dysfunction
				Tatnell, P. J. et al.
				Napsins: new human aspartic proteinases. Distinction between two closely related genes.
				FEBS Lett 441, 43-8 (1998).
				Chuman, Y. et al.
				Napsin A, a member of the aspartic protease family, is abundantly expressed in normal lung and kidney tissue and is expressed in lung adenocarcinomas.
				FEBS Lett 462, 129-34 (1999).

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		711592|Kdap	4.1E-13	[Rattus norvegicus] Kidney-derived aspartic protease-like protein (napsin), aspartic-type endopeptidase; human NAP1 may act as a marker for primary lung adenocarcinoma and lack of human NAP1 in the urine is associated with kidney dysfunction Schauer-Vukasinovic, V. et al. Biochim Biophys Acta 1492, 207-10. (2000).
28	7509113CD1	g182390	1.6E-246	[Homo sapiens] coagulation factor X Messier, T. L. et al. Gene 99, 291-294 (1991)
		344180|F10	1.3E-247	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Extracellular (excluding cell wall)] Coagulation factor X, a vitamin K-dependent serine protease that converts prothrombin (F2) into thrombin, high levels are associated with risk of thrombosis; mutations in corresponding gene cause the coagulation disorder factor X deficiency
				Kaul, R. K. et al.
				Isolation and characterization of human blood-coagulation factor X cDNA.
				Gene 41, 311-4 (1986).
				Racchi, M. et al.
				Human coagulation factor X deficiency caused by a mutant signal peptide that blocks cleavage by signal peptidase but not targeting and translocation to the endoplasmic reticulum.
				J Biol Chem 268, 5735-40 (1993).
		582809|F10	1.6E-187	[Mus musculus] [Hydrolase; Protease (other than proteasomal)] [Extracellular (excluding cell wall)] Coagulation factor X, a vitamin K-dependent serine protease that converts prothrombin (F2) into thrombin; high levels of human F10 predict risk of thrombosis, mutations in the human F10 gene cause the coagulation disorder Factor X deficiency
				Liang, Z. et al.
				Cloning and characterization of a cDNA encoding murine coagulation factor X.

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
				Thromb Haemost 80, 87-91 (1998).
29	7509140CD1	g182831	8.1E-210	[Homo sapiens] coagulation factor X Leytus, S. P. et al. Biochemistry 25, 5098-5102 (1986)
		344180|F10	6.8E-211	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Extracellular (excluding cell wall)] Coagulation factor X, a vitamin K-dependent serine protease that converts prothrombin (F2) into thrombin, high levels are associated with risk of thrombosis; mutations in corresponding gene cause the coagulation disorder factor X deficiency
				Kaul, R. K. et al. (supra)
				Racchi, M. et al. (supra)
		582809|F10	4.8E-149	[Mus musculus] [Hydrolase; Protease (other than proteasomal)] [Extracellular (excluding cell wall)] Coagulation factor X, a vitamin K-dependent serine protease that converts prothrombin (F2) into thrombin; high levels of human F10 predict risk of thrombosis, mutations in the human F10 gene cause the coagulation disorder Factor X deficiency
				Liang, Z. et al. (supra)
30	7509223CD1	g182831	6.9E-246	[Homo sapiens] coagulation factor X Leytus, S. P. et al. (supra)
		344180|F10	5.8E-247	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Extracellular (excluding cell wall)] Coagulation factor X, a vitamin K-dependent serine protease that converts prothrombin (F2) into thrombin, high levels are associated with risk of thrombosis; mutations in corresponding gene cause the coagulation disorder factor X deficiency
				Kaul, R. K. et al. (supra)
				Racchi, M. et al. (supra)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		582809|F10	3.1E-182	[Mus musculus] [Hydrolase; Protease (other than proteasomal)] [Extracellular (excluding cell wall)] Coagulation factor X, a vitamin K-dependent serine protease that converts prothrombin (F2) into thrombin; high levels of human F10 predict risk of thrombosis, mutations in the human F10 gene cause the coagulation disorder Factor X deficiency
				Liang, Z. et al. (supra)
31	7509272CD1	g14348014	2.0E-173	[Homo sapiens] DNA encoding a transmembrane serine protease (Endotheliase 2-S) protein Madison, E. L. et al. Patent: WO 0136604-A
		742472|MSP	8.2E-173	[Homo sapiens] Mosaic serine protease, contains a serine protease domain and regulatory modules having a protein kinase substrate and a low-density lipoprotein receptor Kim, D. R. et al. Biochim Biophys Acta 1518, 204-9. (2001).
		588075|Tmprss2	1.2E-36	[Mus musculus] [Hydrolase; Protease (other than proteasomal)] [Plasma membrane] Epitheliasin, a serine proteinase that contains transmembrane, LDLRA (LDL receptor class A) and SRCR (scavenger receptor cysteine-rich) domains; human TMPRSS2 is highly expressed in androgen-dependent prostate cancer Jacquinet, E. et al. Eur J Biochem 268, 2687-99. (2001).
32	7509327CD1	g14348014	1.0E-87	[Homo sapiens] DNA encoding a transmembrane serine protease (Endotheliase 2-S) protein Madison, E. L. et al. (supra)
		742472|MSP	8.8E-73	[Homo sapiens] Mosaic serine protease, contains a serine protease domain and regulatory modules having a protein kinase substrate and a low-density lipoprotein receptor
				Kim, D. R. et al. (supra)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
33	7504677CD1	g5410288	5.2E-38	[Homo sapiens] FK506-binding protein FKB23 isoform
		473948|LOC51661	4.4E-39	[Homo sapiens] [Isomerase; Chaperones] Member of the FKBP-type peptidyl-prolyl cis-trans isomerase family, has a region of moderate similarity to a region of human FKBP2, which is an FK506-binding protein that binds the immunosuppressive drugs FK506 and rapamycin
		586231|Fkbp7	1.7E-28	[Mus musculus] [Isomerase; Chaperones; Small molecule-binding protein] [Endoplasmic reticulum; Cytoplasmic] FK506-binding protein 7, an endoplasmic reticulum-resident glycoprotein that contains two EF-hand domains and binds calcium, shares a peptidylprolyl cis-trans isomerase motif with other members of the FK506-binding protein family Nakamura, T. et al. Genomics 54, 89-98 (1998).
34	7504534CD1	g2656141	0.0	[Homo sapiens] UnpEL Gray, D. A. et al. Oncogene 10, 2179-2183 (1995) Frederick, A. et al. Oncogene 16, 153-165 (1998)
		338842|USP4	0.0	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Nuclear; Cytoplasmic] Ubiquitin specific protease 4 (ubiquitous nuclear protein human), efficiently cleaves ubiquitin-proline bonds, expression is elevated in small cell lung carcinomas; corresponding gene maps to a chromosomal region frequently rearranged in tumor cells
				Gray, D. A. et al. (supra)
				Frederick, A. et al. (supra)
		586113|Usp4	0.0	[Mus musculus] [Hydrolase; Protease (other than proteasomal); Small molecule-binding protein] [Nuclear] Ubiquitin specific protease 4 (ubiquitous nuclear protein), efficiently cleaves ubiquitin-proline bonds, transforms cells when overexpressed, may bind to the retinoblastoma gene product; expression of human USP4 is elevated in small cell lung carcinomas
				Gupta, K. et al.
				Oncogene 8, 2307-10 (1993).

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
35	7507771CD1	g7108521	4.7E-213	[Arabidopsis thaliana] ubiquitin-protein ligase 1 Bates, P. W. et al. Plant J. 20, 183-195 (1999)
		256316|Y54G2A.30	1.1E-180	[Caenorhabditis elegans] [Ligase] Putative ortholog of S. cerevisiae TOM1 gene product (Protein required for the G2/M transition)
				Huang1, K. et al.
				A HECT domain ubiquitin ligase closely related to the mammalian protein WWP1 is essential for Caenorhabditis elegans embryogenesis.
				Gene 252, 137-145 (2000).
36	7504732CD1	g49645	5.1E-11	[Mesocricetus auratus] P5 Chaudhuri, M. M. et al. Biochem. J. 281 (Pt 3), 645-650 (1992)
		342684|P5	7.1E-12	[Homo sapiens] [Isomerase; Chaperones] Protein disulfide isomerase-related protein, member of the protein disulfide isomerase thioredoxin-containing family of endoplasmic reticulum proteins Hayano, T. et al. Gene 164, 377-8 (1995).
		328208|Rn.2685	8.7E-12	[Rattus norvegicus] [Isomerase; Chaperones] Protein disulfide isomerase-related protein, member of the protein disulfide isomerase thioredoxin-containing family of endoplasmic reticulum proteins, catalyzes disulfide isomerization and the formation of disulfide bonds during protein folding Kramer, B. et al. Biochem J 357, 83-95. (2001).
37	950917CD1	g1164924	1.4E-46	[Mustela vison] plasminogen activator inhibitor type 1 Chuang, T. H. et al. Gene 162, 303-308 (1995)
		581465|Serpine2	6.3E-47	[Mus musculus] [Ligand; Inhibitor or repressor] [Extracellular (excluding cell wall)] Nexin I, a secreted member of the serpin family of serine protease inhibitors, interacts with thrombin and is synthesized in seminal vesicles
				Hagglund, A. C. et al.

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
				Coordinated and cell-specific induction of both physiological plasminogen activators creates functionally redundant mechanisms for plasmin formation during ovulation.
				Endocrinology 137, 5671-7 (1996).
		347042|SERPINE2	8.1E-47	[Homo sapiens] [Inhibitor or repressor] Nexin, a member of the serpin family of serine protease inhibitors that is localized to neuromuscular junctions and to synapses in discrete regions of the brain, regulates alpha-thrombin
				Mbebi, C. et al.
				Protease nexin I expression is up-regulated in human skeletal muscle by injury-related factors.
				J Cell Physiol 179, 305-14 (1999).
		618378|SERPINE1	2.1E-46	[Homo sapiens] [Structural protein; Inhibitor or repressor] Plasminogen activator inhibitor 1, a member of the serpin family of serine proteases and inhibitors, plays a role in regulating blood coagulation by inhibiting fibrinolysis, contributes to tumor progression and is a risk factor for cardiovascular diseases
				Eriksson, P. et al.
				Allele-specific increase in basal transcription of the plasminogen- activator inhibitor 1 gene is associated with myocardial infarction.
				Proc Natl Acad Sci U S A 92, 1851-5 (1995).
38	7459720CD1	g14715064	0.0	[Homo sapiens] proprotein convertase subtilisin/kexin type 7
		341028|PCSK7	0.0	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] Lymphoma proprotein convertase, a member of the proprotein convertase family, a Ca2+-dependent serine protease prohormone convertase that may be involved in cell-cell signaling Meerabux, J. et al. Cancer Res 56, 448-51 (1996).
				Cheng, M. et al.
				Pro-protein convertase gene expression in human breast cancer.

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
				Int J Cancer 71, 966-71 (1997).
		609805|Pcsk7	0.0	[Rattus norvegicus] [Hydrolase; Protease (other than proteasomal)] [Plasma membrane] Lymphoma proprotein convertase, a member of the kexin-like protease family, a subtilisin-kexin like proprotein convertase that associates with the trans-Golgi network transport vesicle
				Seidah, N. G. et al.
				cDNA structure, tissue distribution, and chromosomal localization of rat PC7, a novel mammalian proprotein convertase closest to yeast kexin-like proteinases.
				Proc Natl Acad Sci U S A 93, 3388-93 (1996).
39	7503300CD1	g14602871	1.4E-183	[Homo sapiens] mitochondrial intermediate peptidase
		343090|MIPEP	4.9E-182	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Cytoplasmic; Mitochondrial] Mitochondrial intermediate peptidase, a zinc-dependent metallopeptidase that probably cleaves octapeptide sequences from imported proteins Chew, A. et al. Genomics 40, 493-6 (1997).
		704932|Mipep	2.0E-139	[Rattus norvegicus] [Hydrolase; Protease (other than proteasomal)] [Cytoplasmic; Mitochondrial] Mitochondrial intermediate peptidase, removes amino-terminal octapeptides from nuclear encoded mitochondrial precursor proteins
				Kalousek, F. et al.
				Rat liver mitochondrial intermediate peptidase (MIP): purification and initial characterization.
				EMBO Journal 11, 2803-9. (1992).
40	7503334CD1	g6942155	7.8E-157	[Drosophila melanogaster] putative cytoplasmic aminopeptidase
		691122|FLJ11583	3.9E-221	[Homo sapiens] Protein containing a cytosol aminopeptidase family catalytic domain, has high similarity to uncharacterized C. elegans ZK353.6

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		276703|ZK353.6	1.1E-95	[Caenorhabditis elegans] Member of the cytosol aminopeptidase protein family
				Joshua, G. W.
				Functional analysis of leucine aminopeptidase in Caenorhabditis elegans.
				Mol Biochem Parasitol 113, 223-32. (2001).
41	7503341CD1	g4580549	6.2E-101	[Homo sapiens] LON protease
		340840|PRSS15	5.4E-89	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Cytoplasmic; Mitochondrial] Protease serine 15, an ATP-dependent mitochondrial peptidase that binds single-stranded DNA and may play a role in the regulation of mitochondrial DNA replication and gene expression
				Wang, N. et al.
				A human mitochondrial ATP-dependent protease that is highly homologous to bacterial Lon protease.
				Proc Natl Acad Sci U S A 90, 11247-51 (1993).
				Fu, G. K. et al.
				The human LON protease binds to mitochondrial promoters in a single-stranded, site-specific, strand-specific manner.
				Biochemistry 37, 1905-9. (1998).
		240467|C34B2.6	2.7E-15	[Caenorhabditis elegans] [Protease (other than proteasomal)] [Mitochondrial] Member of the PIM1 serine protease protein family
42	7509936CD1	g14336701	0.0	[Homo sapiens] small optic lobes homolog
				Daniels, R. J. et al.
				Sequence, structure and pathology of the fully annotated terminal 2 Mb of the short arm of human chromosome 16
				Hum. Mol. Genet. 10, 339-352 (2001)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		342796|SOLH	0.0	[Homo sapiens] [Hydrolase; Transcription factor; DNA-binding protein; Protease (other than proteasomal)] Small optic lobes (Drosophila) homolog, a member of the calpain and zinc-finger gene families, has five N-terminal zinc finger-like motifs and a calpain-like protease domain; corresponding gene is a candidate for hereditary cataracts with microphthalmia
				Kamei, M. et al.
				SOLH, a human homologue of the Drosophila melanogaster small optic lobes gene is a member of the calpain and zinc-finger gene families and maps to human chromosome 16p13.3 near CATM (cataract with microphthalmia).
				Genomics 51, 197-206 (1998).
		477348|Solh	0.0	[Mus musculus] [Hydrolase; Protease (other than proteasomal)] Small optic lobes homolog (Drosophila), has five N-terminal zinc finger-like motifs and a calpain-like protease domain, olfactory bulb expression suggests roles in the development and function of sensory neurons, olfaction, vision, and synaptogenesis
				Kamei, M. et al.
				Solh, the mouse homologue of the Drosophila melanogaster small optic lobes gene: organization, chromosomal mapping, and localization of gene product to the olfactory bulb.
				Genomics 64, 82-9 (2000).
43	7509986CD1	g3288916	0.0	[Homo sapiens] aortic carboxypeptidase-like protein ACLP
				Layne, M. D. et al.
				Aortic carboxypeptidase-like protein, a novel protein with discoidin and carboxypeptidase-like domains, is Up-regulated during vascular smooth muscle cell differentiation

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
				J. Biol. Chem. 273, 15654-15660 (1998)
		339938|AEBP1	0.0	[Homo sapiens] [Hydrolase; Inhibitor or repressor; Transcription factor; DNA-binding protein; Protease (other than proteasomal)] [Cytoplasmic] Adipocyte enhancer (AE)-binding protein 1 (aortic carboxypeptidase-like protein), a transcriptional repressor with carboxypeptidase activity, may play a role in adipogenesis
				Ohno, I. et al.
				A cDNA cloning of human AEBP1 from primary cultured osteoblasts and its expression in a differentiating osteoblastic cell line.
				Biochem Biophys Res Commun 228, 411-4 (1996).
		584201|Aebp1	0.0	[Mus musculus] [Hydrolase; Inhibitor or repressor; DNA-binding protein; Transcription factor; Protease (other than proteasomal)] [Nuclear; Cytoplasmic] Adipocyte enhancer(AE)-binding protein 1 (aortic carboxypeptidase-like protein), a transcriptional repressor with carboxypeptidase activity, may play a role in adipogenesis and vascular smooth muscle differentiation, may regulate MAP kinase activity
				Kim, S. W. et al.
				Regulation of adipogenesis by a transcriptional repressor that modulates MAPK activation.
				J Biol Chem 276, 10199-206. (2001).
44	7510010CD1	g183142	0.0	[Homo sapiens] gamma-glutmyl transpeptidase-related protein
				Heisterkamp, N. et al.
				Identification of a human gamma-glutamyl cleaving enzyme related to, but distinct from, gamma-glutamyl transpeptidase
				Proc. Natl. Acad. Sci. U.S.A. 88, 6303-6307 (1991)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		340614|GGTLA1	0.0	[Homo sapiens] [Transferase] Gamma-glutamyl leukotrienase (gamma-glutamyl transpeptidase-related enzyme), converts leukotriene C4 to leukotriene D4 by catalyzing the transfer of the gamma-glutamyl moiety of glutathione
				Heisterkamp, N. et al.
				Identification of a human gamma-glutamyl cleaving enzyme related to, but distinct from, gamma-glutamyl transpeptidase.
				Proc Natl Acad Sci U S A 88, 6303-7 (1991).
		587461|Ggtla1	1.8E-234	[Mus musculus] [Transferase] Gamma-glutamyl leukotrienase (gamma-glutamyl transpeptidase-related enzyme), converts leukotriene C4 to leukotriene D4 by catalyzing the transfer of the gamma-glutamyl moiety of glutathione
				Carter, B. Z. et al.
				Gamma-glutamyl leukotrienase, a gamma-glutamyl transpeptidase gene family member, is expressed primarily in spleen.
				J Biol Chem 273, 28277-85 (1998).
45	7510056CD1	g3288916	7.9E-140	[Homo sapiens] aortic carboxypeptidase-like protein ACLP
				Layne, M. D. et al. (supra)
		339938|AEBP1	6.4E-141	[Homo sapiens] [Hydrolase; Inhibitor or repressor; Transcription factor; DNA-binding protein; Protease (other than proteasomal)] [Cytoplasmic] Adipocyte enhancer (AE)-binding protein 1 (aortic carboxypeptidase-like protein), a transcriptional repressor with carboxypeptidase activity, may play a role in adipogenesis
				Ohno, I. et al. (supra)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		584201|Aebp1	6.8E-114	[Mus musculus] [Hydrolase; Inhibitor or repressor; DNA-binding protein; Transcription factor; Protease (other than proteasomal)] [Nuclear; Cytoplasmic] Adipocyte enhancer (AE)-binding protein 1 (aortic carboxypeptidase-like protein), a transcriptional repressor with carboxypeptidase activity, may play a role in adipogenesis and vascular smooth muscle differentiation, may regulate MAP kinase activity
				Kim, S. W. et al. (supra)
46	7510398CD1	g14348014	3.9E-209	[Homo sapiens] DNA encoding a transmembrane serine protease (Endotheliase 2-S) protein
				Madison, E. L. et al. (supra)
		742472|MSP	5.5E-185	[Homo sapiens] Mosaic serine protease, contains a serine protease domain and regulatory modules having a protein kinase substrate and a low-density lipoprotein receptor
				Kim, D. R. et al. (supra)
		588075|Tmprss2	1.6E-34	[Mus musculus] [Hydrolase; Protease (other than proteasomal)] [Plasma membrane] Epitheliasin, a serine proteinase that contains transmembrane, LDLRA (LDL receptor class A) and SRCR (scavenger receptor cysteine-rich) domains; human TMPRSS2 is highly expressed in androgen-dependent prostate cancer
				Jacquinet, E. et al. (supra)
47	7510498CD1	g3288916	1.5E-280	[Homo sapiens] aortic carboxypeptidase-like protein ACLP
				Layne, M. D. et al. (supra)
		339938|AEBP1	1.2E-281	[Homo sapiens] [Hydrolase; Inhibitor or repressor; Transcription factor; DNA-binding protein; Protease (other than proteasomal)] [Cytoplasmic] Adipocyte enhancer (AE)-binding protein 1 (aortic carboxypeptidase-like protein), a transcriptional repressor with carboxypeptidase activity, may play a role in adipogenesis
				Ohno, I. et al. (supra)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		584201|Aebp1	9.8E-209	[Mus musculus] [Hydrolase; Inhibitor or repressor; DNA-binding protein; Transcription factor; Protease (other than proteasomal)] [Nuclear; Cytoplasmic] Adipocyte enhancer (AE)-binding protein 1 (aortic carboxypeptidase-like protein), a transcriptional repressor with carboxypeptidase activity, may play a role in adipogenesis and vascular smooth muscle differentiation, may regulate MAP kinase activity
				Kim, S. W. et al. (supra)
48	7510044CD1	g9992878	0.0	[Homo sapiens] nicastrin
				Yu, G. et al.
				Nicastrin modulates presenilin-mediated notch/glp-1 signal transduction and beta APP processing
				Nature 407, 48-54 (2000)
49	7504509CD1	g12653735	6.8E-17	[Homo sapiens] aspartyl aminopeptidase
		428674|DNPEP	3.9E-12	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Cytoplasmic] Cytosolic aspartyl aminopeptidase, a member of the M18 family of metalloproteases, has a substrate preference for N-terminal aspartyl and glutamyl residues and may be involved in intracellular peptide metabolism
				Wilk, S. et al.
				Purification, characterization, and cloning of a cytosolic aspartyl aminopeptidase.
				J Biol Chem 273, 15961-70 (1998).
50	7506825CD1	g12804329	6.0E-12	[Homo sapiens] aminoacylase 1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		334030|ACY1	4.9E-13	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Cytoplasmic] Aminoacylase 1, homodimeric zinc-binding enzyme that catalyzes the hydrolysis of N-alpha acylated amino acids; mutational inactivation or dysregulation may contribute to the development of small cell lung cancer
				Cook, R. M. et al.
				Human aminoacylase-1. Cloning, sequence, and expression analysis of a chromosome 3p21 gene inactivated in small cell lung cancer.
				J. Biol. Chem. 268, 17010-7 (1993).
				Cook, R. M. et al.
				Mutational inactivation of aminoacylase-I in a small cell lung cancer cell line.
				Genes Chromosomes Cancer 21, 320-5. (1998).
51	7506828CD1	g178071	9.4E-187	[Homo sapiens] aminoacylase-1
				Miller, Y. E. et al.
				Human aminoacylase-1: cloning, regional assignment to distal chromosome 3p21.1, and identification of a cross-hybridizing sequence on chromosome 18
				Genomics 8, 149-154 (1990)
		334030|ACY1	7.7E-188	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Cytoplasmic] Aminoacylase 1, homodimeric zinc-binding enzyme that catalyzes the hydrolysis of N-alpha acylated amino acids; mutational inactivation or dysregulation may contribute to the development of small cell lung cancer
				Cook, R. M. et al. (1993) (supra)
				Cook, R. M. et al. (1998) (supra)
		713993|C06A6.4	6.8E-84	[Caenorhabditis elegans] [Hydrolase; Protease (other than proteasomal)] [Lysosome/vacuole] Member of the carboxypeptidase yscS protein family

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
52	7506831CD1	g285903	3.0E-203	[Homo sapiens] aminoacylase-1
				Mitta, M. et al.
				The nucleotide sequence of human aminoacylase-1
				Biochim. Biophys. Acta 1174, 201-203 (1993)
		334030|ACY1	2.5E-204	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Cytoplasmic] Aminoacylase 1, homodimeric zinc-binding enzyme that catalyzes the hydrolysis of N-alpha acylated amino acids; mutational inactivation or dysregulation may contribute to the development of small cell lung cancer
				Cook, R. M. et al. (1993) (supra)
				Cook, R. M. et al. (1998) (supra)
53	7509968CD1	g6063386	4.4E-91	[Homo sapiens] kallikrein-like protein 4 KLK-L4
				Yousef, G. M. et al.
				Identification and characterization of KLK-L4, a new kallikrein-like gene that appears to be down-regulated in breast cancer tissues
				J. Biol. Chem. 275, 11891-11898 (2000)
				Diamandis, E. P. et al.
				The new kallikrein gene family: implications in carcinogenesis
				Trends Endocrinol. Metab. 11, 54-60 (2000)
		625969|KLK13	3.6E-92	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] Kallikrein-like protein, a serine protease that is down-regulated in breast cancer, may be involved in breast cancer pathogenesis and/or progression
				Clements, J. et al.
				The expanded human kallikrein (KLK) gene family: genomic organisation, tissue-specific expression and potential functions.
				Biol Chem 382, 5-14. (2001).

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		609150|Prss20	7.9E-33	[Mus musculus] [Hydrolase; Protease (other than proteasomal)] [Extracellular (excluding cell wall)] Hippostasin, a secreted protein in the trypsin family of serine proteases
				Mitsui, S. et al.
				cDNA cloning and tissue-specific splicing variants of mouse hippostasin/TLSP (PRSS20)(1).
				Biochim Biophys Acta 1494, 206-210 (2000).
54	7510232CD1	g1401352	3.7E-149	[Homo sapiens] apoptotic cysteine protease Mch5 isoform alpha
				Fernandes-Alnemri, T. et al.
				CPP32, a novel human apoptotic protein with homology to Caenorhabditis elegans cell death protein Ced-3 and mammalian interleukin-1 beta-converting enzyme
				J. Biol. Chem. 269, 30761-30764 (1994)
55	7510233CD1	g12862693	3.5E-265	[Homo sapiens] caspase-8
				Hadano, S. et al.
				Cloning and Characterization of Three Novel Genes, ALS2CR1, ALS2CR2, and ALS2CR3, in the Juvenile Amyotrophic Lateral Sclerosis (ALS2) Critical Region at Chromosome 2q33-q34: Candidate Genes for ALS2
				Genomics 71, 200-213 (2001)
56	7510304CD1	g306886	6.2E-211	[Homo sapiens] hepsin (serine protease) precursor
				Leytus, S. P. et al.
				A novel trypsin-like serine protease (hepsin) with a putative transmembrane domain expressed by human liver and hepatoma cells
				Biochemistry 27, 1067-1074 (1988)
		335864|HPN	5.1E-212	[Homo sapiens] [Hydrolase; Protease (other than proteasomal)] [Plasma membrane] Hepsin, a transmembrane serine protease implicated in cell growth control and initiation of blood coagulation; overexpressed in ovarian cancer

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
				Tanimoto, H. et al.
				Hepsin, a cell surface serine protease identified in hepatoma cells, is overexpressed in ovarian cancer.
				Cancer Res 57, 2884-7. (1997).
				Luo, J. et al.
				Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling.
				Cancer Res 61, 4683-8. (2001).
		583247|Hpn	2.1E-190	[Mus musculus] [Hydrolase; Protease (other than proteasomal)] Hepsin, a transmembrane serine protease implicated in cell growth control and initiation of blood coagulation; human HPN is overexpressed in ovarian cancer
				Vu, T. K. H. et al.
				Identification and cloning of the membrane-associated serine protease, hepsin, from mouse preimplantation embryos.
				J Biol Chem 272, 31315-20 (1997).
				Kawamura, S. et al.
				Complete nucleotide sequence, origin of isoform and functional characterization of the mouse hepsin gene.
				Eur J Biochem 262, 755-64 (1999).
				Wu, Q. et al.
				Generation and characterization of mice deficient in hepsin, a hepatic transmembrane serine protease.
				J Clin Invest 101, 321-6 (1998).
57	7510461CD1	g6942153	5.2E-119	[Drosophila melanogaster] putative cytoplasmic aminopeptidase
58	7510392CD1	g12653457	2.1E-59	[Homo sapiens] Similar to stromal cell-derived factor 2

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		742542|SDF2	1.7E-60	[Homo sapiens] [Transferase] [Extracellular (excluding cell wall)] Stromal cell-derived factor 2, a putative secreted protein with similarity to S. cerevisiae Pmt1p and Pmt2p, which are dolichyl phosphate-D-mannose:protein mannosyltransferases
				Hamada, T. et al.
				Isolation and characterization of a novel secretory protein, stromal cell-derived factor-2 (SDF-2) using the signal sequence trap method.
				Gene 176, 211-4 (1996).
				Fukuda, S. et al.
				Murine and Human SDF2L1 Is an Endoplasmic Reticulum Stress-Inducible Gene and Encodes a New Member of the Pmt/rt Protein Family.
				Biochem Biophys Res Commun 280, 407-414. (2001).
		581301|Sdf2	9.9E-58	[Mus musculus] [Extracellular (excluding cell wall)] Stromal cell-derived factor 2, a putative secreted protein with similarity to S. cerevisiae Pmt1p and Pmt2p, which are dolichyl phosphate-D-mannose:protein mannosyltransferases
				Hamada, T. et al. (supra)
				Fukuda, S. et al. (supra)

Table 3

<table>
<tr><th>SEQ ID NO:</th><th>Incyte Polypeptide ID</th><th>Amino Acid Residues</th><th>Potential Phosphorylation Sites</th><th>Potential Glycosylation Sites</th><th>Signature Sequences, Domains and Motifs</th><th>Analytical Methods and Databases</th></tr>
<tr><td>1</td><td>7313196CD1</td><td>321</td><td>S26 S197 S243 T73 T126 T130 T252</td><td></td><td>signal_cleavage: M1-C19</td><td>SPSCAN</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M1-C19, M1-S20</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Serpin (serine protease inhibitor): M88-P321, P24-D87</td><td>HMMER_PFAM</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Serpins proteins
BL00284: N50-T73, T103-L144, V215-F241, D297-P321</td><td>BLIMPS_BLOCKS</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Serpins signature: D273-P321</td><td>PROFILESCAN</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>SERPIN INHIBITOR PROTEASE SERINE SIGNAL PRECURSOR GLYCOPROTEIN PLASMA PROTEIN PROTEINASE
PD000192: L27-N86, K74-P321</td><td>BLAST_PRODOM</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>SERPINS
DM00112|P07093|24-396: P24-K114, D87-K320
DM00112|I48717|25-395: L25-V149, M88-K320
DM00112|P13909|30-400: L30-D87, M88-I318
DM00112|P20961|29-400: S32-D87, M88-I318</td><td>BLAST_DOMO</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Serpins signature: F294-I304</td><td>MOTIFS</td></tr>
</table>

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
2	6465289CD1	1309	S88 S91 S123 S128 S145 S204 S226 S322 S413 S417 S435 S450 S458 S490 S498 S688 S757 S769 S817 S845 S876 S883 S913 S924 S935 S972 S973 S993 S997 S1002 S1005 S1068 S1074 S1083 S1135 S1184 S1209 S1230 S1231 S1264 T35 T108 T225 T303 T390 T399 T452 T508 T586 T613 T679 T713 T721 T739 T760 T801 T857 T985 T988 T995 T1041 T1091 T1131 T1162 Y296 Y377 Y497 Y558 Y622 Y647 Y748 Y1011	N65 N161 N719 N1078 N1181	Ubiquitin carboxyl-terminal hydrolases family: V168-W199	HMMER_PFAM
					Ubiquitin carboxyl-terminal hydrolase family: N463-L543	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Ubiquitin carboxyl-terminal hydrolases family 2 proteins BL00972: G169-L186, W248-C257, I277-C291, I466-S490, D493-K514	BLIMPS_BLOCKS
					UBIQUITIN CARBOXYL TERMINAL HYDROLASE 64E EC 3.1.2.15 THIOLESTERASE UBIQUITIN-SPECIFIC PROCESSING PROTEASE DE-UBIQUITINATING ENZYME CONJUGATION THIOL NUCLEAR PROTEIN PD143046: I507-R715, H909-A957	BLAST_PRODOM
					PROTEASE UBIQUITIN HYDROLASE UBIQUITIN-SPECIFIC ENZYME DE-UBIQUITINATING CARBOXYL TERMINAL THIOLESTERASE PROCESSING CONJUGATION PD017412: R300-Q410	BLAST_PRODOM
					UBIQUITIN CARBOXYL-TERMINAL HYDROLASES FAMILY 2 DM00659|P50101|209-458: N172-D402, D428-G477 DM00659|Q09879|217-457: N172-E412, Y467-G477 DM00659|P40818|782-1103: F261-I401, K462-T515 DM00659|P38187|434-611: T271-D402	BLAST_DOMO
					Ubiquitin carboxyl-terminal hydrolases family 2 signature 1: G169-Q184	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Ubiquitin carboxyl-terminal hydrolases family 2 signature 2: Y467-Y484	MOTIFS
3	7506357CD1	987	S79 S86 S212 S280 S377 S404 S413 S530 S606 S639 S746 S801 S817 T60 T167 T172 T193 T261 T332 T465 T600 T887 Y15	N476	ThiF family: R431-P576, Q31-G163	HMMER_PFAM
					Repeat in ubiquitin-activating (UBA) pro: D786-Y929	HMMER_PFAM
					Ubiquitin-activating enzyme proteins BL00536: A588-N631, G775-Q805, N806-Y844, S413-V454, G462-P505, I523-E561	BLIMPS_BLOCKS
					Ubiquitin-activating enzyme signatures: A354-L402	PROFILESCAN
					UBIQUITIN-ACTIVATING ENZYME E1 UBIQUITIN CONJUGATION LIGASE REPEAT MULTIGENE FAMILY PROTEIN PD005434: C160-L430	BLAST_PRODOM
					UBIQUITIN-ACTIVATING ENZYME E1 UBIQUITIN CONJUGATION LIGASE REPEAT MULTIGENE FAMILY PROTEIN PD005366: H577-T698, F692-V774, F757-I828	BLAST_PRODOM
					ENZYME UBIQUITIN-ACTIVATING UBIQUITIN CONJUGATION LIGASE E1 REPEAT PROTEIN MULTIGENE FAMILY PD003325: V578-F623, P780-P852	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					PROTEIN ENZYME UBIQUITIN-ACTIVATING UBIQUITIN E1 CONJUGATION LIGASE REPEAT BIOSYNTHESIS MULTIGENE PD000731: Q31-F159, L436-P576	BLAST_PRODOM
					UBIQUITIN-ACTIVATING ENZYME DM02861|P41226|628-1010: T629-K696, F681-L987	BLAST_DOMO
					UBIQUITIN-ACTIVATING ENZYME DM00412|P41226|4-255: L4-V256 DM00412|P41226|402-626: P403-E628	BLAST_DOMO
					UBIQUITIN-ACTIVATING ENZYME DM02915|P41226|257-400: K257-L402	BLAST_DOMO
					Ubiquitin-activating enzyme active site: P597-P605	MOTIFS
4	6878857CD1	227	S130 S147 S186		signal_cleavage: M1-A15	SPSCAN
					Signal Peptide: M1-A16, M1-F18	HMMER
					Trypsin: I24-I219	HMMER_PFAM
					Trypsin-like serine protease: K23-I219	
					Serine proteases, trypsin family, active sites: I159-Q202	PROFILESCAN
					Chymotrypsin serine protease family (S1) signature PR00722: T83-A97, K173-V185	BLIMPS_PRINTS
					PROTEASE SERINE PRECURSOR SIGNAL HYDROLASE ZYMOGEN GLYCOPROTEIN FAMILY MULTIGENE FACTOR PD000046: I24-A71, S43-Y155, I91-I219	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					TRYPSIN DM00018|P07478|24-242: I24-Y45, G26-I223 DM00018|P35030|81-299: I24-G44, G26-I223 DM00018|S37538|36-254: K23-G44, G26-I223 DM00018|P07477|24-242: I24-Y45, G26-I223	BLAST_DOMO
					Serine proteases, trypsin family, serine active site: D174-V185	MOTIFS
5	7506021CD1	727	S18 S93 S107 S122 S126 S176 S300 S339 S363 S388 S405 S416 S423 S430 S447 S465 S488 S492 S521 S538 S655 S660 S705 T278 T390 T626 Y130	N88 N349 N401 N681	Ulp1 protease family, C-terminal catalytic domain: F563-R724	HMMER_PFAM
					PROTEASE SENTRIN/SUMO-SPECIFIC SENTRIN-SPECIFIC SMT4 CHROMOSOME C17A5.07C 4930538C18 RIK-RELATED MOLECULE PD009801: K602-S664, V673-L720	BLAST_PRODOM
6	7503356CD1	143	S18 S22 S126 T128	N124	Eukaryotic thiol (cysteine) proteases active sites: Q63-S126	PROFILESCAN
					Pathogenesis-related proteins Bet v I family signature: G41-Q95	PROFILESCAN
					Calpain cysteine protease (C2) family signature PR00704: Q95-A111, Q28-A51, W71-L93	BLIMPS_PRINTS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					PROTEASE CALPAIN HYDROLASE SUBUNIT NEUTRAL THIOL LARGE CALCIUMACTIVATED PROTEINASE CANP PD001545: L43-W138	BLAST_PRODOM
					CALPAIN CATALYTIC DOMAIN DM01305|P17655|1-505: Q28-W138 DM01305|P00789|3-507: Q28-W141 DM01305|P07384|11-517: Q28-W141 DM01305|S57196|12-574: Y17-W138	BLAST_DOMO
					Eukaryotic thiol (cysteine) proteases cysteine active site: Q95-A106	MOTIFS
7	7509052CD1	31	T29		Cyclophilin type peptidyl-prolyl cis-trans isomerase: T5-E23	HMMER_PFAM
					PROTEIN PEX17MER1 INTERGENIC REGION PD143519: N3-L27	BLAST_PRODOM
					PRECURSOR CELL ADHESION GLYCOPROTEIN TRANSMEMBRANE CALCIUMBINDING REPEAT SIGNAL PHOSPHORYLATION DESMOCOLLIN PD003960: G14-E23	BLAST_PRODOM
					POLYPROTEIN PROTEASE GENOME CONTAINS: COAT PROTEIN VP4 P1A VP2 P1B PD037652: V2-G18, G14-V24	BLAST_PRODOM
					REPLICATION PROTEIN C PLASMID REPC PD040010: V2-T29	BLAST_PRODOM
8	7503366CD1	630	S47 S76 S109 S113 S137 S205 S228 S248 S280 S348 S369 S454 S480	N283 N311 N374	Ubiquitin carboxyl-terminal hydrolases family: V162-Y193	HMMER_PFAM
					Ubiquitin carboxyl-terminal hydrolases family 2 proteins BL00972: G163-L180, E251-T260	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			S495 S504 S505 S518 S521 S551 T130 T134 T207 T260 T485 T490 T533 T535		PROTEASE UBIQUITIN HYDROLASE ENZYME UBIQUITINSPECIFIC CARBOXYLTERMINAL DEUBIQUITINATING THIOLESTERASE PROCESSING CONJUGATION PD000590: P161-N287	BLAST_PRODOM
9	7505933CD1	354	S43 S115 S155 S179 S200 S219 S243 S257 S289 S314 T68 T127 T228 T259		Ubiquitin interaction motif: M258-A275, P187-Q204	HMMER_PFAM
					PROTEASOME 26S REGULATORY SUBUNIT S5A PROTEIN MULTIUBIQUITIN CHAIN BINDING ANTISECRETORY PD009810: M1-K129, D119-E169	BLAST_PRODOM
					PROTEASOME 26S REGULATORY SUBUNIT S5A MULTIUBIQUITIN CHAIN BINDING PROTEIN ANTISECRETORY PD015781: E231-L336 PD117215: E182-S233	BLAST_PRODOM
10	7507064CD1	132	S16 S117 T25 T44 T85 Y33		Pepsin (A1) aspartic protease family signature PR00792: S22-T35, W105-D120	BLIMPS_PRINTS
					PROTEASE ASPARTYL HYDROLASE PRECURSOR SIGNAL ZYMOGEN GLYCOPROTEIN ASPARTIC PROTEINASE MULTIGENE PD000182: F55-A131, M1-R46	BLAST_PRODOM
					EUKARYOTIC AND VIRAL ASPARTYL PROTEASES DM00126|P03954|16-386: G53-A129, L5-T44 DM00126|P00794|18-379: M53-A131, M1-A68 DM00126|P16476|16-381: G53-A131, L5-T44 DM00126|P28713|16-385: G53-A131, L5-T44	BLAST_DOMO
11	1439986CD1	316	S40 S119 S147		DnaJ domain: D4-G68	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			S162 T156 T174 T251 T285 Y171		DnaJ C terminal region: T192-L314	HMMER_PFAM
					DnaJ molecular chaperone homology domain: Q3-K60	HMMER_PFAM
					Nt-dnaJ domain proteins BL00636: Q19-K35, F45-D65	BLIMPS_BLOCKS
					dnaJ domains signatures and profile: R25-P85	PROFILESCAN
					DnaJ protein family signature PR00625: S15-L34, F45-D65, I182-K198, N227-C244	BLIMPS_PRINTS
					PROTEIN CHAPERONE DNAJ HEAT SHOCK DNA REPLICATION REPEAT ANTIGEN T PD000231: D4-G68	BLAST_PRODOM
					PROTEIN HEAT SHOCK CHAPERONE DNAJ HOMOLOG HSP403 DROJ1 F54D5.8 PSI PD015361: F80-S172	BLAST_PRODOM
					NT-DNAJ DOMAIN DM00098|P25685|1-107: M1-G106 DM00098|S23509|1-108: Y5-G105 DM00098|P25686|1-108: Y5-G105 DM00156|P25685|198-272: K176-V249	BLAST_DOMO
					Nt-dnaJ domain signature: F45-Y64	MOTIFS
12	2008979CD1	531	S15 S125 S140 S142 S176 S322 S335 S370 S383 S423 S430 S436 S523 T136 T160 T325 Y81	N355 N434	DnaJ domain: C3-R69	HMMER_PFAM
					DnaJ molecular chaperone homology domai: K2-E61	HMMER_SMART
					zinc finger: L314-H338, I482-H506	HMMER_SMART
					U1-like zinc finger: Y311-L345	HMMER_SMART
					Zinc finger, C2H2 type: L314-H338, I482-H506	HMMER_PFAM
					Zinc finger, C2H2 type, domain proteins BL00028: C316-H332	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Nt-dnaJ domain proteins BL00636: E18-K34, F46-D66	BLIMPS_BLOCKS
					DNAJ PROTEIN FAMILY SIGNATURE PR00625: A14-D33, F46-D66	
					dnaJ domains signatures and profile: R24-S85	PROFILESCAN
					PROTEIN CHAPERONE NUCLEAR DNABINDING ZINCFINGER METALBINDING ZUOTIN C6B12.08 CHROMOSOME I PD009324: T109-V217	BLAST_PRODOM
					PROTEIN COILED COIL CHAIN MYOSIN REPEAT HEAVY ATPBINDING FILAMENT HEPTAD PD000002: E180-K392, R182-Q395, K181-D398, K167-K390, K215-E433	BLAST_PRODOM
					PROTEIN CHAPERONE DNAJ HEAT SHOCK DNA REPLICATION REPEAT ANTIGEN T PD000231: H4-D66	BLAST_PRODOM
					PROTEIN REPEAT TROPOMYOSIN COILED COIL ALTERNATIVE SPLICING SIGNAL PRECURSOR CHAIN PD000023: K215-Q395	BLAST_PRODOM
					NT-DNAJ DOMAIN DM00098|P53863|1-107: M1-L73 DM00098|P35514|1-116: H4-G79 DM00098|P30725|1-102: K2-S85 DM00098|P25685|1-107: K2-G75	BLAST_DOMO
					Cytochrome c family heme-binding site signature: C484-S489	MOTIFS
					Nt-dnaJ domain signature: F46-Y65	MOTIFS
					EF-hand calcium-binding domain: D281-L293	MOTIFS

Table 3

<table>
<tr><th>SEQ ID NO:</th><th>Incyte Polypeptide ID</th><th>Amino Acid Residues</th><th>Potential Phosphorylation Sites</th><th>Potential Glycosylation Sites</th><th>Signature Sequences, Domains and Motifs</th><th>Analytical Methods and Databases</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Zinc finger, C2H2 type, domain: C316-H338, C484-H506</td><td>MOTIFS</td></tr>
<tr><td>13</td><td>90073157CD1</td><td>108</td><td>T69 Y94</td><td></td><td>signal_cleavage: M1-G38</td><td>SPSCAN</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M1-G38</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>DINACETYLCHITOBIASE PRECURSOR HYDROLASE GLYCOSIDASE SIGNAL LYSOSOME GLYCOPROTEIN PD034983: D40-L108</td><td>BLAST_PRODOM</td></tr>
<tr><td>14</td><td>7506782CD1</td><td>345</td><td>S80 S193 S217 S275 S338 T20 T40 T68 T98 T159</td><td>N157 N256 N336</td><td>signal_cleavage: M1-G18</td><td>SPSCAN</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M1-A16, M1-G18, M1-F21, M1-P22, M1-A25</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Serpin (serine protease inhibitor): Y174-P332, A19-N160</td><td>HMMER_PFAM</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Serine protease inhibitors:V31-P332</td><td>HMMER_PFAM</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Serpins proteins BL00284: N45-T68, V147-S167, V225-F251, D308-P332</td><td>BLIMPS_BLOCKS</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Serpins signature: S284-T334</td><td>PROFILESCAN</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>SERPIN INHIBITOR PROTEASE SERINE SIGNAL PRECURSOR GLYCOPROTEIN PLASMA PROTEIN PROTEINASE PD000192:Q175-N336, E24-A170</td><td>BLAST_PRODOM</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>SERPINS
DM00112|I48717|25-395:F163-V329, E23-T159, V11-P50
DM00112|A53120|45-416:V176-V329, N32-E169
DM00112|P01014|2-386:Q175-R328, A27-D165
DM00112|P07093|24-396: Y174-V329, E23-D165</td><td>BLAST_DOMO</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Serpins signature: V305-I315</td><td>MOTIFS</td></tr>
</table>

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
15	7506941CD1	493	S12 S67 S105 S141 S192 S249 S250 S461 T81 T161 T213 T256 T265 T277 T387 Y217 Y455		Dipeptidyl peptidase IV PF00930: Q62-M88, I160-R170, D214-T259, K367-P394, R419-L439	BLIMPS_PFAM
					DIPEPTIDYL IV HYDROLASE PROTEASE SERINE PEPTIDASE DIPEPTIDASE TRANSMEMBRANE GLYCOPROTEIN PROTEIN PD003048:E361-E481 PD003086:P153-V323, R69-T81	BLAST_PRODOM
					PROLYL ENDOPEPTIDASE FAMILY SERINE DM02461|P18962|229-817:F364-R463, I160-M368, S67-P78 DM02461|P33894|340-930:F364-I493, K155-M368, H106-D149, R69-K79 DM02461|P27487|192-765:G342-E481, F150-M368, Y70-F82, D3-D18 DM02461|P42659|335-862:A374-E481, A159-K359, A65-I98	BLAST_DOMO
16	7507072CD1	204	S14 S124 S147 S199 T141	N36 N171	signal_cleavage: M1-G33	SPSCAN
					Carboxypeptidase activation peptide: Q41-R118	HMMER_PFAM
					Zinc carboxypeptidase: Y139-K178	HMMER_PFAM
					Zinc carboxypeptidases, zinc-binding region 1 proteins BL00132: Y139-E179	BLIMPS_BLOCKS
					CARBOXYPEPTIDASE PRECURSOR HYDROLASE ZINC ZYMOGEN SIGNAL B A 3DSTRUCTURE A1 PD005637:G33-E117	BLAST_PRODOM
					ZINC CARBOXYPEPTIDASES, ZINC-BINDING REGION 1 DM00683|P15085|112-418:R129-L203	BLAST_DOMO

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
17	7507083CD1	224	S53 S80 S90 S117 S167 S186 T21 T214 Y110 Y133	N115 N140	ICE-like protease (caspase) p20 domain: N105-Q144	HMMER_PFAM
					Interleukin-1B converting enzyme signature PR00376: R114-G132, G132-F150	BLIMPS_PRINTS
					ICEC_MOUSE // CASPASE12 PRECURSOR EC 3.4.22. HYDROLASE THIOL PROTEASE APOPTOSIS ZYMOGEN PD103766:V12-D91	BLAST_PRODOM
18	7509097CD1	277	T184 S15 S127 S132 T184		Signal Peptide: M27-A60	HMMER
					Cytosolic domain: W215-D277 Transmembrane domain: L192-F214 Non-cytosolic domain: M1-Q191	TMHMMER
					signal_cleavage: M1-A50	SPSCAN
					LDL-receptor class A (LDLRA) domain proteins BL01209: C237-E249	BLIMPS_BLOCKS
19	7509118CD1	449	S15 S127 S132 S283 S285 S332 T184 T320 T349 T380 T428	N281 N318	Signal Peptide: M27-A60	HMMER
					Trypsin: I352-F394	HMMER_PFAM
					Cytosolic domain: M1-Q191 Transmembrane domain: L192-F214 Non-cytosolic domain: W215-L449	TMHMMER
					Serine proteases, trypsin family, histidine proteins BL00134: C377-C393	BLIMPS_BLOCKS
					LDL-receptor class A (LDLRA) domain proteins BL01209: C237-E249	BLIMPS_BLOCKS
					Serine proteases, trypsin family, active sites: L369-K419	PROFILESCAN
					Low density lipoprotein (LDL) receptor signature PR00261: E228-E249, E228-E249, E228-E249, E228-E249, E228-E249, E228-E249	BLIMPS_PRINTS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Chymotrypsin serine protease family (S1) signature PR00722: G378-C393	BLIMPS_PRINTS
					V8 serine protease family signature PR00839: C377-F394	BLIMPS_PRINTS
					TRANSMEMBRANE PROTEASE, SERINE 2 EC 3.4.21. HYDROLASE PROTEASE SIGNALANCHOR PD072395: P117-R351	BLAST_PRODOM
					TRYPSIN DM00018|P05981|163-403: I352-G410 DM00018|P06868|553-782: T349-N413 DM00018|I48685|32-230: I352-C393 DM00018|P20918|576-808: G350-Y415	BLAST_DOMO
					Serine proteases, trypsin family, histidine active site: L388-C393	MOTIFS
20	7509312CD1	451	S61 S66 S217 S219 S266 S356 T118 T254 T283 T314 T367 T393 T430	N215 N252 N365	signal_cleavage: M1-G62	SPSCAN
					Signal Peptide: M1-A34	HMMER
					Trypsin: I286-H403	HMMER_PFAM
					Cytosolic domain: M1-Q125 Transmembrane domain: L126-F148 Non-cytosolic domain: W149-L451	TMHMMER
					Serine proteases, trypsin family, histidine proteins BL00134: C311-C327	BLIMPS_BLOCKS
					LDL-receptor class A (LDLRA) domain proteins BL01209: C171-E183	BLIMPS_BLOCKS
					Serine proteases, trypsin family, active sites: W299-A354	PROFILESCAN
					Low density lipoprotein (LDL) receptor signature PR00261: E162-E183, E162-E183, E162-E183, E162-E183, E162-E183, E162-E183	BLIMPS_PRINTS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Chymotrypsin serine protease family (S1) signature PR00722: G312-C327, E370-L384	BLIMPS_PRINTS
					TRANSMEMBRANE PROTEASE, SERINE 2 EC 3.4.21. HYDROLASE PROTEASE SIGNALANCHOR PD072395: P51-R285	BLAST_PRODOM
					PROTEASE SERINE PRECURSOR SIGNAL HYDROLASE ZYMOGEN GLYCOPROTEIN FAMILY MULTIGENE FACTOR PD000046: I286-S428	BLAST_PRODOM
					TRYPSIN DM00018|P26262|391-624: I286-P394 DM00018|P03952|392-624: V287-P394 DM00018|A57014|45-284: I286-P401 DM00018|P14272|391-624: I286-P394	BLAST_DOMO
					Serine proteases, trypsin family, histidine active site: L322-C327	MOTIFS
21	90126902CD1	485	S66 S69 S469 T17 T119 T291 T401 T479	N197 N218	signal_cleavage: M1-G39	SPSCAN
					Signal Peptide: M1-G44, M1-A42	HMMER
					Thrombospondin type 1 domain: T49-C96	HMMER_PFAM
					Thrombospondin type 1 repeats: W48-P97	HMMER_SMART
					PROTEIN PROCOLLAGEN THROMBOSPONDIN MOTIFS NPROTEINASE A DISINTEGRIN METALLOPROTEASE WITH ADAMTS1 PD011654:Q134-C200	BLAST_PRODOM
					Myb DNA-binding domain repeat signature 1: W468-I476	MOTIFS
22	7509352CD1	258	S39 S85 S107 T29 T56 T123 T175	N131	signal_cleavage: M1-P24	SPSCAN

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Signal Peptide: M1-P19, M1-P21, M1-P24	HMMER
					Lysyl oxidase: V57-S238	HMMER_PFAM
					Scavenger receptor cysteine-rich domain: P37-H124	HMMER_PFAM
					Scavenger receptor Cys-rich: L32-S133	HMMER_SMART
					Lysyl oxidase copper-binding region proteins BL00926: Q90-D126, D126-R156, Y158-F197, Q198-S238	BLIMPS_BLOCKS
					Speract receptor repeated domain signature: L16-W96	PROFILESCAN
					Lysyl oxidase signature PR00074: I95-T123, T123-C150, T152-I180, D181-E209, S210-N237	BLIMPS_PRINTS
					Speract receptor signature PR00258: G52-A63, A67-E77, D98-C112	BLIMPS_PRINTS
					LYSYL OXIDASE PROTEIN LYSINE PRECURSOR SIGNAL 6-OXIDASE OXIDOREDUCTASE COPPER GLYCOPROTEIN HOMOLOG PD012364: C112-C233	BLAST_PRODOM
					ANTIGEN PRECURSOR SIGNAL M130 TRANSMEMBRANE GLYCOPROTEIN REPEAT VARIANT CYTOPLASMIC PROTEIN PD000767: V35-H124	BLAST_PRODOM
					OXIDASE; LYSINE; LYSYL; COPPER; DM04978|Q05063|1-419: C112-C233 DM04978|A48501|238-573: C112-C233 DM04978|A47005|16-410: C112-C233	BLAST_DOMO
					SPERACT RECEPTOR AMINO-TERMINAL DM00148|P21757|345-450: T30-C112	BLAST_DOMO

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
23	7509341CD1	252	S99 S202 T124 T210 T232 Y89	N173	signal_cleavage: M1-W21	SPSCAN
					Signal Peptide: M1-W21, M1-C28, M1-P25	HMMER
					Trypsin: I37-W161, G162-I207	HMMER_PFAM
					Trypsin-like serine protease: R36-I207	HMMER_SMART
					Kringle domain proteins BL00021: C63-F80, I145-G166, G166-I207	BLIMPS_BLOCKS
					Serine proteases, trypsin family, histidine proteins BL00134: C63-C79, T158-I181, P194-I207	BLIMPS_BLOCKS
					Type I fibronectin domain proteins BL01253: C63-A76, G157-C170, W176-T210	BLIMPS_BLOCKS
					Serine proteases, trypsin family, active sites: W55-P103	PROFILESCAN
					Serine proteases, trypsin family, active sites: L137-R190	PROFILESCAN
					Chymotrypsin serine protease family (S1) signature PR00722: G64-C79, T123-L137, G157-V169	BLIMPS_PRINTS
					PROTEASE SERINE PRECURSOR SIGNAL HYDROLASE ZYMOGEN GLYCOPROTEIN FAMILY MULTIGENE FACTOR PD000046: I37-C160, C160-I207	BLAST_PRODOM
					TRYPSIN DM00018|Q02844|29-268: I37-N173, C160-I207 DM00018|P14272|391-624: I37-W161, C160-I207 DM00018|P26262|391-624: I37-W161, G157-I207 DM00018|P03952|392-624: V38-W161, C160-I207	BLAST_DOMO
					Leucine zipper pattern: L67-L88	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Serine proteases, trypsin family, histidine active site: L74-C79	MOTIFS
					Serine proteases, trypsin family, serine active site: T158-V169	MOTIFS
24	7509367CD1	102			signal_cleavage: M1-W21	SPSCAN
					Signal Peptide: M1-W21, M1-C28, M1-P25	HMMER
					Kringle domain proteins BL00021: C63-F80	BLIMPS_BLOCKS
					Serine proteases, trypsin family, histidine proteins BL00134: C63-C79	BLIMPS_BLOCKS
					Serine proteases, trypsin family, active sites: W55-A101	PROFILESCAN
					Chymotrypsin serine protease family (S1) signature PR00722: G64-C79	BLIMPS_PRINTS
					TRYPSIN DM00018|P15944|31-270: I37-C79 DM00018|P05981|163-403: I37-F80 DM00018|I48685|32-230: I37-C79 DM00018|P21845|31-271: I37-C79	BLAST_DOMO
					Serine proteases, trypsin family, histidine active site: L74-C79	MOTIFS
25	7500455CD1	62	S16 S32 S47 S51 S57 T33 T46		UBIQUITIN CARBOXYL TERMINAL HYDROLASE 11 EC 3.1.2.15 THIOLESTERASE UBIQUITIN-SPECIFIC PROCESSING PROTEASE 13 DE-UBIQUITINATING ENZYME KIAA0055 CONJUGATION THIOL MULTIGENE FAMILY PD142716: M1-I37	BLAST_PRODOM
26	7510401CD1	181	S9 S87 T5 T65 T95		Ubiquitin associated domain: E3-E41	HMMER_SMART

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					UBA/TS-N domain: A2-H42	HMMER_PFAM
					PUTATIVE GLIALBLASTOMA CELL DIFFERENTIATION-RELATED PROTEIN PD179110: A6-E108	BLAST_PRODOM
27	7504702CD1	104	S60		signal_cleavage: M1-P21	SPSCAN
					Signal Peptide: M1-P21, M1-G23, M1-A24, M1-L26	HMMER
28	7509113CD1	444	S43 S63 S74 S139 S146 S330 S345 S394 S421 T111 T155 T179 T249 T289 T311 T365 Y126 Y340	N177 N187	signal_cleavage: M1-G21	SPSCAN
					Signal Peptide: M1-L19, M1-L20, M1-G21, M1-S23, M1-F25, M1-R28	HMMER
					EGF-like domain: C90-C121	HMMER_PFAM
					Vitamin K-dependent carboxylation/gamma-carboxyglutamic (GLA) domain: L45-D86	HMMER_PFAM
					Trypsin: I191-I418	HMMER_PFAM
					Cytosolic domain: M1-H6 Transmembrane domain: L7-I26 Non-cytosolic domain: R27-K444	TMHMMER
					Serine proteases, trypsin family, histidine proteins BL00134: C217-C233, D369-I392, Y405-I418	BLIMPS_BLOCKS
					Type I fibronectin domain proteins BL01253: G115-P158, C217-A230, H324-G362, E368-T381, Y387-S421	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Vitamin K-dependent carboxylation domain: L24-D103	PROFILESCAN
					Serine proteases, trypsin family, active sites: L208-G255	PROFILESCAN
					Serine proteases, trypsin family, active sites: I354-R401	PROFILESCAN
					Coagulation factor GLA domain signature PR00001: F44-C57, M58-F71, E72-D86	BLIMPS_PRINTS
					Type II EGF-like signature PR00010: D86-N97, Q98-L105, G106-F116, E117-L123	BLIMPS_PRINTS
					Chymotrypsin serine protease family (S1) signature PR00722: G218-C233, T274-I288, E368-V380	BLIMPS_PRINTS
					PROTEASE SERINE PRECURSOR SIGNAL HYDROLASE ZYMOGEN GLYCOPROTEIN FAMILY MULTIGENE FACTOR PD000046: I191-T331, N248-I418	BLAST_PRODOM
					FACTOR X GLYCOPROTEIN COAGULATION PRECURSOR SIGNAL HYDROLASE STUART SERINE PROTEASE PD016023: P125-R190	BLAST_PRODOM
					TRYPSIN DM00018|P00742|234-465: R190-M422 DM00018|S49075|231-463: R190-M422 DM00018|P00743|233-464: R190-M422 DM00018|P25155|240-471: R190-I418	BLAST_DOMO
					EGF-like domain signature 1: C110-C121	MOTIFS
					EGF-like domain signature 2: C110-C121	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Calcium-binding EGF-like domain pattern signature: D86-C110	MOTIFS
					Vitamin K-dependent carboxylation domain: F44-W81	MOTIFS
					Serine proteases, trypsin family, histidine active site: L228-C233	MOTIFS
					Serine proteases, trypsin family, serine active site: D369-V380	MOTIFS
29	7509140CD1	377	S43 S63 S74 S183 S190 S263 S278 S327 S354 T111 T125 T156 T199 T223 T298 Y273	N221 N231	signal_cleavage: M1-G21	SPSCAN
					Signal Peptide: M1-L19, M1-L20, M1-G21, M1-S23, M1-F25, M1-R28	HMMER
					EGF-like domain: C90-C121, C129-C164	HMMER_PFAM
					Vitamin K-dependent carboxylation/gamma-carboxyglutamic (GLA) domain: L45-D86	HMMER_PFAM
					Trypsin: N257-I351, I235-L252	HMMER_PFAM
					Cytosolic domain: M1-H6 Transmembrane domain: L7-I26 Non-cytosolic domain: R27-K377	TMHMMER
					Apple domain proteins BL00495: A294-W328, G329-T357	BLIMPS_BLOCKS
					Type I fibronectin domain proteins BL01253: G115-A158, N257-G295, E301-T314, Y320-S354	BLIMPS_BLOCKS
					Vitamin K-dependent carboxylation domain: L24-K102	PROFILESCAN

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Serine proteases, trypsin family, active sites: I287-R334	PROFILESCAN
					Coagulation factor GLA domain signature PR00001: F44-C57, M58-F71, E72-D86	BLIMPS_PRINTS
					Type II EGF-like signature PR00010: D86-N97, Q98-L105, G106-F116, E117-L123	BLIMPS_PRINTS
					Chymotrypsin serine protease family (S1) signature PR00722: E301-V313	BLIMPS_PRINTS
					FACTOR X GLYCOPROTEIN COAGULATION PRECURSOR SIGNAL HYDROLASE STUART SERINE PROTEASE PD016023: T167-R234	BLAST_PRODOM
					PROTEASE SERINE PRECURSOR SIGNAL HYDROLASE ZYMOGEN GLYCOPROTEIN FAMILY MULTIGENE FACTOR PD000046: I235-E258, L266-I351	BLAST_PRODOM
					GAMMA-CARBOXYGLUTAMIC ACID VITAMIN K GLYCOPROTEIN PRECURSOR CALCIUM-BINDING PROTEIN COAGULATION PLASMA PD001826: L45-Y84	BLAST_PRODOM
					TRYPSIN DM00018|P00742|234-465: R234-E258, N257-M355 DM00018|S49075|231-463: R234-G260, N257-M355 DM00018|P00743|233-464: R234-E258, N257-M355	BLAST_DOMO

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					COAGULATION FACTOR X DM02035|P00742|130-232: S130-T233	BLAST_DOMO
					Cell attachment sequence: R227-D229	MOTIFS
					Aspartic acid and asparagine hydroxylation site: C101-C112	MOTIFS
					EGF-like domain signature 1: C110-C121	MOTIFS
					EGF-like domain signature 2: C110-C121, C149-C164	MOTIFS
					Calcium-binding EGF-like domain pattern signature: D86-C110	MOTIFS
					Vitamin K-dependent carboxylation domain: F44-W81	MOTIFS
					Serine proteases, trypsin family, serine active site: D302-V313	MOTIFS
30	7509223CD1	442	S43 S63 S74 S183 S190 S328 S343 S392 S419 T111 T125 T156 T199 T223 T287 T309 T363 Y338	N221 N231	signal_cleavage: M1-G21	SPSCAN
					Signal Peptide: M1-L19, M1-L20, M1-G21, M1-S23, M1-F25, M1-R28	HMMER
					EGF-like domain: C90-C121, C129-C164	HMMER_PFAM
					Vitamin K-dependent carboxylation/gamma-carboxyglutamic (GLA) domain: L45-D86	HMMER_PFAM
					Trypsin: E258-I416, I235-L252	HMMER_PFAM
					Epidermal growth factor-like domain: Q89-E122, L128-I165	HMMER_SMART

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Calcium-binding EGF-like domain: D86-E122, N133-I165	HMMER_SMART
					Domain containing Gla (gamma-carboxyglutamic): E22-K85	HMMER_SMART
					Trypsin-like serine protease: R234-I416	HMMER_SMART
					Cytosolic domain: M1-H6 Transmembrane domain: L7-I26 Non-cytosolic domain: R27-K442	TMHMMER
					Apple domain proteins BL00495: P247-P285, A359-W393, G394-T422	BLIMPS_BLOCKS
					Type I fibronectin domain proteins BL01253: G115-A158, H322-G360, E366-T379, Y385-S419	BLIMPS_BLOCKS
					Vitamin K-dependent carboxylation domain: L24-D103	PROFILESCAN
					Serine proteases, trypsin family, active sites: I352-R399	PROFILESCAN
					Coagulation factor GLA domain signature PR00001: F44-C57, M58-F71, E72-D86	BLIMPS_PRINTS
					Type II EGF-like signature PR00010: D86-N97, Q98-L105, G106-F116, E117-L123	BLIMPS_PRINTS
					Chymotrypsin serine protease family (S1) signature PR00722: T272-I286, E366-V378	BLIMPS_PRINTS
					PROTEASE SERINE PRECURSOR SIGNAL HYDROLASE ZYMOGEN GLYCOPROTEIN FAMILY MULTIGENE FACTOR PD000046: I235-V259, E260-I416	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					FACTOR X GLYCOPROTEIN COAGULATION PRECURSOR SIGNAL HYDROLASE STUART SERINE PROTEASE PD016023: T167-R234	BLAST_PRODOM
					TRYPSIN DM00018|P00742|234-465: R234-E258, N257-M420 DM00018|S49075|231-463: R234-E256, N257-M420 DM00018|P00743|233-464: R234-E258, N254-M420 DM00018|P25155|240-471: G87-C110, R234-E258, E260-I416	BLAST_DOMO
					Cell attachment sequence: R227-D229	MOTIFS
					Aspartic acid and asparagine hydroxylation site: C101-C112	MOTIFS
					EGF-like domain signature 1: C110-C121	MOTIFS
					EGF-like domain signature 2: C110-C121, C149-C164	MOTIFS
					Calcium-binding EGF-like domain pattern signature: D86-C110	MOTIFS
					Vitamin K-dependent carboxylation domain: F44-W81	MOTIFS
					Serine proteases, trypsin family, serine active site: D367-V378	MOTIFS
31	7509272CD1	375	S217 S219 S266 S356 T118 T254 T283 T314 T367	N215 N252 N365	Signal Peptide: M1-A34	HMMER
					Trypsin: I286-Y366	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Cytosolic domain: M1-Q125 Transmembrane domain: L126-F148 Non-cytosolic domain: W149-S375	TMHMMER
					LDL-receptor class A (LDLRA) domain proteins BL01209: C171-E183	BLIMPS_BLOCKS
					Serine proteases, trypsin family, histidine proteins BL00134: C311-C327	BLIMPS_BLOCKS
					Serine proteases, trypsin family, active sites: W299-A354	PROFILESCAN
					Chymotrypsin serine protease family (S1) signature PR00722: G312-C327	BLIMPS_PRINTS
					TRANSMEMBRANE PROTEASE, SERINE 2 EC 3.4.21. HYDROLASE PROTEASE SIGNAL ANCHOR PD072395: S35-R285	BLAST_PRODOM
					TRYPSIN DM00018|P05981|163-403: I286-G344 DM00018|P03952|392-624: V287-E370 DM00018|P26262|391-624: I286-E370 DM00018|P03951|389-621: V287-D374	BLAST_DOMO
					Serine proteases, trypsin family, histidine active site: L322-C327	MOTIFS
32	7509327CD1	204	S101 S106 T158		signal_cleavage: M1-A24	SPSCAN
					Signal Peptide: M1-A34	HMMER
					Cytosolic domain: M1-Q165 Transmembrane domain: L166-T188 Non-cytosolic domain: G189-S204	TMHMMER
33	7504677CD1	86	S31 S50 S74 T49	N45	signal_cleavage: M1-A23	SPSCAN
					Signal Peptide: M1-R25, M1-A23, M5-A23	HMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					FKBP-type peptidyl-prolyl cis-trans isomerase: E44-R75	HMMER_PFAM
					Cytosolic domain: M1-H6 Transmembrane domain: F7-Q24 Non-cytosolic domain: R25-F86	TMHMMER
					Phosphoenolpyruvate carboxylase active sites: F15-G68	PROFILESCAN
					FK506-BINDING PROTEIN PD059301: M5-C46	BLAST_PRODOM
34	7504534CD1	882	S76 S157 S196 S244 S364 S642 S675 S680 S829 S859 T15 T82 T201 T570 T693 T746 T774 T775 T786 T790 T809 T840 Y620	N240 N443 N650 N850 N857	Ubiquitin carboxyl-terminal hydrolases family: C302-D333	HMMER_PFAM
					Ubiquitin carboxyl-terminal hydrolases family 2 proteins BL00972: G303-L320, Y389-L398, I447-C461, V864-Y883	BLIMPS_BLOCKS
					PROTEASE UBIQUITIN HYDROLASE UBIQUITIN-SPECIFIC CARBOXYL TERMINAL THIOLESTERASE PROCESSING DE-UBIQUITINATING ENZYME UBIQUITOUS PD009843: M1-E79, D52-F283	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					PROTEASE UBIQUITIN-SPECIFIC UBIQUITIN CARBOXYL TERMINAL HYDROLASE THIOLESTERASE PROCESSING DE-UBIQUITINATING ENZYME UBIQUITOUS PD152487: R587-K770	BLAST_PRODOM
					PROTEASE UBIQUITIN HYDROLASE ENZYME UBIQUITIN-SPECIFIC CARBOXYL TERMINAL DE-UBIQUITINATING THIOLESTERASE PROCESSING CONJUGATION PD000590: P299-T470	BLAST_PRODOM
					PROTEASE UBIQUITIN HYDROLASE UBIQUITIN-SPECIFIC CARBOXYL TERMINAL THIOLESTERASE PROCESSING DE-UBIQUITINATING ENZYME PROTEIN PD011543: F471-E586	BLAST_PRODOM
					UBIQUITIN CARBOXYL TERMINAL HYDROLASES FAMILY 2 DM00659|P35123|139-432: L307-L600 DM00659|P51784|41-331: L307-R587 DM00659|P40818|782-1103: L307-L481, Q754-Y882	BLAST_DOMO
					UBIQUITIN; HYDROLASE; TERMINAL; CARBOXYL; DM08763|P35123|433-705: Y601-G875	BLAST_DOMO
					Ubiquitin carboxyl-terminal hydrolases family 2 signature 1: G303-Q318	MOTIFS
					Ubiquitin carboxyl-terminal hydrolases family 2 signature 2: Y865-Y882	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
35	7507771CD1	4374	S114 S142 S153 S235 S327 S345 S539 S545 S575 S648 S649 S661 S726 S727 S853 S962 S990 S1056 S1298 S1311 S1343 S1368 S1370 S1395 S1599 S1600 S1624 S1636 S1648 S1736 S1800 S1907 S1999 S2029 S2034 S2144 S2270 S2275 S2338 S2372 S2377 S2424 S2431 S2463 S2501 S2546 S2619 S2662 S2749 S2850 S2855 S2927 S2936 S2952 S2953 S2963 S3072 S3167 S3234 S3255 S3256 S3262 S3363 S3401 S3407 S3447 S3565 S3592 S3609 S3682 S3695 S3797 S3802 S3808 S3818 S3852 S3929 S4030 S4038 S4084 S4140 S4264 T11 T217 T249 T316 T340 T440 T484 T493	N74 N98 N483 N742 N747 N783 N784 N1341 N1634 N1635 N1806 N2091 N2273 N2364 N2389 N2727 N2733 N3002 N3049 N3201 N3399	Domain Homologous to E6-AP Carboxyl Terminal: R4036-A4374	HMMER_SMART

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			T672 T785 T909 T1028 T1123 T1214 T1245 T1304 T1438 T1503 T1509 T1579 T1659 T1703 T1729 T1731 T1735 T1778 T1833 T1905 T1941 T2035 T2170 T2185 T2201 T2309 T2416 T2536 T2549 T2705 T2821 T2866 T2889 T2951 T3054 T3131 T3202 T3242 T3246 T3247 T3312 T3359 T3450 T3549 T3673 T3714 T3719 T3737 T3747 T3892 T3952 T3961 T4144 T4243 T4303 T4332 Y1791 Y4271		Ubiquitin associated domain: N1318-T1354	HMMER_SMART

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					HECT-domain (ubiquitin-transferase): M4067-A4374	HMMER_PFAM
					UBA/TS-N domain: V1317-H1355	HMMER_PFAM
					WWE domain: Y1609-L1680	HMMER_PFAM
					HECT-domain (ubiquitin-transferase) PF00632: R4187-G4193, W4280-P4307, L4335-E4366	BLIMPS_PFAM
					PROTEIN LIGASE UBIQUITIN CONJUGATION REPEAT UBIQUITIN PROTEIN DNA BINDING PROBABLE ONCOGENIC PD002225: L1113-H1151, R4063-G4182, L4104-G4372	BLAST_PRODOM
					UBIQUITIN PROTEIN LIGASE SYSTEM F17O7.14 D8035.1P PD025703: E32-R693 G2454-S2477	BLAST_PRODOM
					LIGASE UBIQUITIN-PROTEIN RELATED SYSTEM F17O7.15 MEMBRANE INNER Y67D8C.G TOM1 F14J16.10 PD147211: E2955-T3143, A3791-L3992, L3449-T3500, L3977-V4019	BLAST_PRODOM
					UBIQUITIN LIGASE SYSTEM PROTEIN D8035.1P PD044151: E2349-S2501, P2894-R3043, G3156-E3241, A3540-T3640, P3831-I3911, F3955-L3992, G3013-E3075, T2649-E2731, G3417-T3501, E2298-P2472, E2276-D2353 N2405-L2467	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					HECT DOMAIN DM01690|P51593|9-306: N4075-L4373 DM01690|P39940|513-808: N4075-G4372 DM01690|A38919|785-1082: F4074-A4374 DM01690|P53119|615-909: N4075-I4364	BLAST_DOMO
					Leucine zipper pattern: L652-L673, L834-L855, L1052-L1073, L3172-L3193, L3179-L3200, L3846-L3867	MOTIFS
					ATP/GTP-binding site motif A (P-loop): A1029-S1036	MOTIFS
36	7504732CD1	89	S34		signal_cleavage: M1-C24	SPSCAN
					Signal Peptide: M1-C24, M1-D30	HMMER
					Thioredoxin: C24-T85	HMMER_PFAM
					Protein disulfide-isomerase domain: D32-N89	HMMER_TIGRFAM
					Thioredoxin family proteins BL00194: L45-K57	BLIMPS_BLOCKS
					Thioredoxin family active site: E29-P76	PROFILESCAN
					Thioredoxin family signature PR00421: W44-W52, W52-P61	BLIMPS_PRINTS
					THIOREDOXIN FAMILY DM00054|P38660|160-264: K25-T85 DM00054|P34329|187-293: N40-T85 DM00054|P13667|171-277: E33-S86 DM00054|P38657|375-479: E29-S86	BLAST_DOMO
					Thioredoxin family active site: L45-W63	MOTIFS
37	950917CD1	424	S137 S203 S271 S288 S333 S408 T83 T107 T197	N46 N136 N148	signal_cleavage: M1-R17	SPSCAN
					Signal Peptide: M1-A18, M1-G20	HMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					SERine Proteinase INhibitors: L35-I402	HMMER_SMART
					Serpin (serine protease inhibitor): R23-I402	HMMER_PFAM
					Serpins proteins BL00284: N49-T72, A184-M225, V296-F322, D378-I402	BLIMPS_BLOCKS
					Serpins signature: G355-I399	PROFILESCAN
					SERPIN INHIBITOR PROTEASE SERINE SIGNAL PRECURSOR GLYCOPROTEIN PLASMA PROTEIN PROTEINASE PD000192: T31-I399	BLAST_PRODOM
					SERPINS DM00112|P13909|30-400: L29-I399 DM00112|I48717|25-395: L29-I399 DM00112|P07093|24-396: L29-I399 DM00112|P20961|29-400: T31-I399	BLAST_DOMO
					Serpins signature: F375-L385	MOTIFS
38	7459720CD1	791	S285 S382 S508 S519 S547 S624 S640 S695 S716 S733 S771 T177 T193 T394 T433 T446 T582 T593 T670 T724 T749 Y664	N167 N175 N241 N517 N769	signal_cleavage: M1-G37	SPSCAN
					Proprotein convertase P-domain: V488-W623	HMMER_PFAM
					Subtilase family: F145-W476	HMMER_PFAM
					Cytosolic domain: M1-D12 Transmembrane domain: A13-V35 Non-cytosolic domain: M36-C791	TMHMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Serine proteases, subtilase family, aspartic acid proteins BL00136: V185-I197, N226-A238, G410-G420	BLIMPS_BLOCKS
					Serine proteases, subtilase family, active sites: R164-P216	PROFILESCAN
					Serine proteases, subtilase family, active sites: D208-D262	PROFILESCAN
					Serine proteases, subtilase family, active sites: S382-V442	PROFILESCAN
					Subtilisin serine protease family (S8) signature PR00723: G178-I197, N224-A237, T409-M425	BLIMPS_PRINTS
					PROTEASE SERINE PRECURSOR CONVERTASE SPC7 PC7 SIGNAL HYDROLASE TRANSMEMBRANE PC8 PD023050: D627-C791	BLAST_PRODOM
					PROTEASE PRECURSOR SERINE HYDROLASE SIGNAL GLYCOPROTEIN ZYMOGEN CONVERTASE ENDOPROTEASE PROHORMONE PD000717: A482-S621	BLAST_PRODOM
					PROTEASE SERINE PRECURSOR CONVERTASE SPC7 PC7 SIGNAL HYDROLASE TRANSMEMBRANE PC8 PD021546: M1-S53	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					SERINE PROTEASES, SUBTILASE FAMILY, HISTIDINE DM00108|S35366|351-610: I166-D436 DM00108|P30430|351-610: I166-D436 DM00108|P26016|351-610: I166-D436 DM00108|P29122|183-443: D165-D436	BLAST_DOMO
					Leucine zipper pattern: L759-L780	MOTIFS
					ATP/GTP-binding site motif A (P-loop): G295-T302	MOTIFS
					Serine proteases, subtilase family, histidine active site: H228-A238	MOTIFS
					Serine proteases, subtilase family, serine active site: G410-G420	MOTIFS
39	7503300CD1	352	S37 S54 S72 S244 S258 S287 S302 S325 S329 T147 T154 T179 T332	N241	signal_cleavage: M1-L17	SPSCAN
					HYDROLASE METALLOPROTEASE ZINC OLIGOPEPTIDASE PRECURSOR MITOCHONDRIAL ENDOPEPTIDASE MITOCHONDRION TRANSIT PEPTIDE PD002945: A73-N224 L282-K345	BLAST_PRODOM
					MITOCHONDRIAL INTERMEDIATE PEPTIDASE PRECURSOR HYDROLASE MIP METALLOPROTEASE ZINC TRANSIT PEPTIDE PD043824: G11-Q103	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					ZINC; METALLOPEPTIDASE; NEUTRAL; OLIGOPEPTIDASE; DM01184|Q01992|30-697: A33-S351 DM01184|P37932|40-760: A45-T223, I238-M338 DM01184|P51980|43-766: L58-P226, D260-K345 DM01184|Q10415|32-751: G64-N347	BLAST_DOMO
40	7503334CD1	495	S113 S242 S304 S404 S489 T107 T168 T294 T313 T351	N145 N292 N311	Cytosol aminopeptidase family, catalytic: V150-A457	HMMER_PFAM
					Cytosol aminopeptidase proteins BL00631: D158-E171, I227-A260, A285-V325, I337-G352, P382-C415, V431-H442	BLIMPS_BLOCKS
					Cytosol aminopeptidase signature PR00481: I227-K244, M249-I270, E286-T307, V308-A328, I337-G352	BLIMPS_PRINTS
					AMINOPEPTIDASE HYDROLASE LEUCINE ZINC LAP LEUCYL CYTOSOL PROLINE PROLYL PROTEIN PD002804: F93-L435	BLAST_PRODOM
					CYTOSOL AMINOPEPTIDASE DM01145|P34629|152-482: L143-R461 DM01145|P47631|117-443: V150-E464 DM01145|P30184|183-516: G149-L456 DM01145|P31427|236-570: G133-L456	BLAST_DOMO
					ATP/GTP-binding site motif A (P-loop): G240-T247	MOTIFS
					Cytosol aminopeptidase signature: N312-L319	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
41	7503341CD1	239	S72 S80 S181 S199 S224 T116 T217	N174 N235	signal_cleavage: M1-A48	SPSCAN
					PROTEASE HYDROLASE SERINE ATP-BINDING MITOCHONDRIAL PRECURSOR MITOCHONDRION TRANSIT PEPTIDE LON PD008588: P107-G189	BLAST_PRODOM
					MITOCHONDRIAL LON PROTEASE HOMOLOG PRECURSOR EC 3.4.21. HYDROLASE SERINE ATP-BINDING MITOCHONDRION TRANSIT PEPTIDE PD053058: M23-W55	BLAST_PRODOM
					MITOCHONDRIAL LON PROTEASE HOMOLOG PRECURSOR HYDROLASE SERINE ATP-BINDING MITOCHONDRION TRANSIT PD027209: G56-D84	BLAST_PRODOM
					ATP-DEPENDENT SERINE PROTEASES, LON FAMILY, SERINE DM04994|Q09769|21-227: T116-G189 DM04994|S62421|21-227: T116-G189	BLAST_DOMO

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
42	7509936CD1	936	S24 S42 S100 S173 S247 S291 S296 S308 S317 S337 S456 S487 S684 S706 S720 S747 S835 S872 T3 T54 T307 T328 T350 T465 T538 T819 Y708	N485 N522 N745	Calpain family cysteine protease: F488-H793	HMMER_PFAM
					Zn-finger in Ran binding protein and others: R412-L441, P143-L172, D340-H369, T3-K32, V43-P73	HMMER_PFAM
					Calpain-like thiol protease family.: E469-V801	HMMER_SMART
					Zinc finger domain: G5-P29, G414-P438, G145-P169, Q46-T70, T342-S366	HMMER_SMART
					Calpain cysteine protease (C2) family signature PR00704: A472-P495, Q546-A562, A579-C604, C609-L632, G634-L661, S769-C790	BLIMPS_PRINTS
					SMALL OPTIC LOBES HOMOLOG PD184656: M1-S487	BLAST_PRODOM
					PROTEASE CALPAIN HYDROLASE SUBUNIT NEUTRAL THIOL LARGE CALCIUM-ACTIVATED PROTEINASE CANP PD001545: F488-V792	BLAST_PRODOM
					PROTEIN SMALL OPTIC LOBES W05G11.4 T11A5.6 ALTERNATIVE SPLICING ZINC FINGER HYDROLASE PD017266: S794-P925	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					CALPAIN CATALYTIC DOMAIN DM01305|P27398|943-1453: N16-R30, N156-S173, R453-W901 DM01305|P07384|11-517: W474-R800, H862-F898 DM01305|P00789|3-507: W474-D795, A829-F898 DM01305|A48764|1-507: W474-D795, V880-F898	BLAST_DOMO
					Eukaryotic thiol (cysteine) proteases cysteine active site: Q546-A557	MOTIFS
43	7509986CD1	1136	S43 S113 S166 S261 S348 S392 S465 S566 S593 S916 S941 S1059 T30 T63 T81 T128 T139 T150 T161 T213 T273 T332 T353 T378 T380 T424 T448 T451 T452 T825 T857 T899 T1041 T1063 T1073 T1081 T1089 T1097 T1124 Y249	N480 N528 N1017	signal_cleavage: M1-P25	SPSCAN
					Signal Peptide: M1-A18, M1-G22, M1-P25, M1-T30	HMMER
					Coagulation factor 5/8 C-terminal domain: K384-C540	HMMER_SMART
					F5/8 type C domain: P387-V537	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Zinc carboxypeptidase: H564-E674, G844-H905, R906-Y949, E743-T778	HMMER_PFAM
					Zinc carboxypeptidases, zinc-binding region 1 proteins BL00132: H564-I604, P616-V629, R647-A687, P728-N754	BLIMPS_BLOCKS
					Carboxypeptidase A metalloprotease (M14) family signature PR00765: P616-L630, A760-Y773	BLIMPS_PRINTS
					PROTEIN AEBP1 AORTIC CARBOXYPEPTIDASE-LIKE ACLP CARBOXYPEPTIDASE AE-BINDING TRANSCRIPTIONAL REPRESSOR PD152287: R1000-F1136	BLAST_PRODOM
					AORTIC CARBOXYPEPTIDASE-LIKE PROTEIN ACLP CARBOXYPEPTIDASE PD173393: P201-D313	BLAST_PRODOM
					CARBOXYPEPTIDASE PRECURSOR SIGNAL HYDROLASE ZINC ZYMOGEN PROTEIN D B GP180 CARBOXYPEPTIDASE PD001916: H564-Q678, E743-G918, I908-P940	BLAST_PRODOM
					PROTEIN AEBP1 AORTIC CARBOXYPEPTIDASE-LIKE ACLP CARBOXYPEPTIDASE AE-BINDING TRANSCRIPTIONAL REPRESSOR PD152055: V675-T742	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					ZINC CARBOXYPEPTIDASES, ZINC-BINDING REGION 1 DM00683|S51739|130-497: Y548-T914 DM00683|P15169|15-341: F562-D774, T806-G907 DM00683|P04836|3-333: D558-T778, D796-G907	BLAST_DOMO
					DISCOIDIN I N-TERMINAL DM00516|S51739|1-128: M419-S547	BLAST_DOMO
					Zinc carboxypeptidases, zinc-binding region 1 signature: P616-L638	MOTIFS
					Coagulation factors 5/8 type C domain (FA58C) signature 1: A432-G461	MOTIFS
					Coagulation factors 5/8 type C domain (FA58C) signature 2: P524-C540	MOTIFS
44	7510010CD1	617	S66 S120 S377 S417 S537 T101 T197 T222 T237 T256 T325 T583 Y617	N98 N204 N303 N347 N535 N581	signal_cleavage: M1-A21	SPSCAN
					Signal Peptide: M1-A19, M1-A21, M1-A26, M1-S30	HMMER
					Gamma-glutamyltranspeptidase: R58-E613	HMMER_PFAM
					g_glut_trans: gamma-glutamyltranspeptidase: G37-R609	HMMER_TIGRFAM
					Cytosolic domain: M1-G6 Transmembrane domain: A7-L29 Non-cytosolic domain: S30-Y617	TMHMMER
					Gamma-glutamyltranspeptidase proteins BL00462: A44-M86, I138-H174, F202-T256, T388-E427, P463-I475	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Gamma-glutamyltranspeptidase signature: P445-A498	PROFILESCAN
					TRANSFERASE GAMMA-GLUTAMYLTRANSPEPTIDASE ACYLTRANSFERASE ZYMOGEN GLUTATHIONE BIOSYNTHESIS PRECURSOR GLYCOPROTEIN TRANSMEMBRANE GGT PD002339: I60-I248, G155-E613	BLAST_PRODOM
					GAMMA-GLUTAMYLTRANSPEPTIDASE 5 PRECURSOR EC 2.3.2.2 GAMMA GLUTAMYLTRANSFERASE GGT REL TRANSFERASE ACYLTRANSFERASE ZYMOGEN GLYCOPROTEIN TRANSMEMBRANE GLUTATHIONE BIOSYNTHESIS SIGNAL ANCHOR PD127342: I23-Q64	BLAST_PRODOM
					GAMMA-GLUTAMYLTRANSPEPTIDASE DM01065|P36269|49-585: D49-K539 Q569-Y617 DM01065|P19440|45-568: A48-H523 Q569-Y617 DM01065|JC4570|44-568: D49-H523 Q569-G616 DM01065|A35074|45-367: A48-R340	BLAST_DOMO
					Cell attachment sequence: R368-D370	MOTIFS
					Gamma-glutamyltranspeptidase signature: T388-G412	MOTIFS
45	7510056CD1	316	S34 S121 S165 S238 S282 S288 T105 T126 T151 T153 T197 T221 T224 T225 T313	N253	Coagulation factor 5/8 C-terminal domain: K157-C299	HMMER_SMART

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					F5/8 type C domain: P160-Q289	HMMER_PFAM
					AORTIC CARBOXYPEPTIDASE-LIKE PROTEIN ACLP CARBOXYPEPTIDASE AEBP1 PD061618: M87-P159	BLAST_PRODOM
					AORTIC CARBOXYPEPTIDASE-LIKE PROTEIN ACLP CARBOXYPEPTIDASE PD173393: E21-D86	BLAST_PRODOM
					GLYCOPROTEIN PRECURSOR SIGNAL FACTOR REPEAT PROTEIN NEUROPILIN CELL DOMAIN COAGULATION PD000875: P160-Y261	BLAST_PRODOM
					DISCOIDIN I N-TERMINAL DM00516|S51739|1-128: M192-M268	BLAST_DOMO
					Coagulation factors 5/8 type C domain (FA58C) signature 1: A205-G234	MOTIFS
46	7510398CD1	418	S96 S101 S252 S254 S301 T153 T289 T318 T349 T397	N250 N287	Signal Peptide: M1-A34	HMMER
					Trypsin: I321-F363	HMMER_PFAM
					Cytosolic domain: M1-Q160 Transmembrane domain: L161-F183 Non-cytosolic domain: W184-L418	TMHMMER
					LDL-receptor class A (LDLRA) domain proteins BL01209: C206-E218	BLIMPS_BLOCKS
					Serine proteases, trypsin family, histidine proteins BL00134: C346-C362	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Serine proteases, trypsin family, active sites: L338-K388	PROFILESCAN
					Chymotrypsin serine protease family (S1) signature PR00722: G347-C362	BLIMPS_PRINTS
					TRANSMEMBRANE PROTEASE, SERINE 2 EC 3.4.21. HYDROLASE PROTEASE SIGNAL ANCHOR PD072395: P86-R320	BLAST_PRODOM
					TRYPSIN DM00018|P05981|163-403: I321-G379 DM00018|P06868|553-782: T318-N382 DM00018|I48685|32-230: I321-C362 DM00018|P20918|576-808: G319-Y384	BLAST_DOMO
					Serine proteases, trypsin family, histidine active site: L357-C362	MOTIFS
47	7510498CD1	543	S43 S113 S166 S261 S348 S392 S465 S509 S515 T30 T63 T81 T128 T139 T150 T161 T213 T273 T332 T353 T378 T380 T424 T448 T451 T452 T540 Y249	N480	signal_cleavage: M1-P25	SPSCAN
					Signal Peptide: M1-A18, M1-G22, M1-P25, M1-T30	HMMER
					Coagulation factor 5/8 C-terminal domain: K384-C526	HMMER_SMART

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					F5/8 type C domain: P387-Q516	HMMER_PFAM
					AORTIC CARBOXYPEPTIDASE-LIKE PROTEIN ACLP CARBOXYPEPTIDASE PD173393: P201-D313	BLAST_PRODOM
					AORTIC CARBOXYPEPTIDASE-LIKE PROTEIN ACLP CARBOXYPEPTIDASE AEBP1 PD061618: M314-P386	BLAST_PRODOM
					AORTIC CARBOXYPEPTIDASE-LIKE PROTEIN ACLP CARBOXYPEPTIDASE PD170167: C20-P116	BLAST_PRODOM
					PROTEIN TOPOISOMERASE I DNA ISOMERASE REPEAT DNABINDING INTERMEDIATE FILAMENT HEPTAD PD000422: P25-L417	BLAST_PRODOM
					DISCOIDIN I N-TERMINAL DM00516|S51739|1-128: M419-M495	BLAST_DOMO
					do NEUROFILAMENT; TRIPLET; DM04498|P12036|434-1019: P25-R394	BLAST_DOMO
					H1; HISTONE; DM03514|S62122|104-232: E46-P387	BLAST_DOMO
					NEUROFILAMENT; TRIPLET; DM04498|P19246|716-1085: A18-Q399	BLAST_DOMO
					Coagulation factors 5/8 type C domain (FA58C) signature 1: A432-G461	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
48	7510044CD1	742	S8 S268 S373 S394 S417 S422 S488 S499 S505 S674 T148 T301 T313 T405 T516 T647 T657 T680	N78 N88 N220 N233 N237 N297 N420 N450 N468 N497 N539 N563 N595 N606 N613 N645	Signal Peptide: M1-F34	HMMER
					Cytosolic domain: K726-Y742 Transmembrane domain: L703-A725 Non-cytosolic domain: M1-E702	TMHMMER
					NICASTRIN PRECURSOR SIGNAL GLYCOPROTEIN TRANSMEMBRANE HOMOLOG PD147358: S168-A739	BLAST_PRODOM
					Trp-Asp (WD) repeats signature: V310-V324	MOTIFS
49	7504509CD1	47				
50	7506825CD1	74	T2		signal_cleavage: M1-A63	SPSCAN
51	7506828CD1	343	S79 S189 T2 T41 T201 T214 T250 Y107	N270	Peptidase family M20/M25/M40: F16-A220	HMMER_PFAM
					ArgE / dapE / ACY1 / CPG2 / yscS family proteins BL00758: I139-G150, S192-S208	BLIMPS_BLOCKS
					HYDROLASE PROTEIN DIPEPTIDE PD00794: S79-P85, G202-M211	BLIMPS_PRODOM
					AMINOACYLASE1 PROTEIN N-ACYL-L-AMINO ACID AMIDOHYDROLASE ACY1 HYDROLASE ZINC ACETYLATION SIMILAR C10C5.3 PD011402: V249-V336	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					PROTEIN HYDROLASE DESUCCINYLASE ZINC SUCCINYL-DIAMINOPIMELATE DEACETYLASE CARBOXYPEPTIDASE ACETYL-ORNITHINE BIOSYNTHESIS COBALT PD001449: D30-G132	BLAST_PRODOM
					ArgE / dapE / ACY1 / CPG2 / yscS family signature 1: I75-V84	MOTIFS
					ArgE / dapE / ACY1 / CPG2 / yscS family signature 2: A111-G150	MOTIFS
52	7506831CD1	373	S44 S154 T2 T166 T179 T219 T280 Y72	N217 N300	Peptidase family M20/M25/M40: G32-T312, F16-Y31	HMMER_PFAM
					ArgE / dapE / ACY1 / CPG2 / yscS family proteins BL00758: I104-G115, S157-S173	BLIMPS_BLOCKS
					HYDROLASE PROTEIN DIPEPTIDASE PD00794: S44-P50, G167-M176, N228-R241	BLIMPS_PRODOM
					AMINOACYLASE1 PROTEIN N-ACYL-L-AMINO ACID AMIDOHYDROLASE ACY1 HYDROLASE ZINC ACETYLATION SIMILAR C10C5.3 PD011402: V279-V366	BLAST_PRODOM
					ArgE / dapE / ACY1 / CPG2 / yscS family signature 1: I40-V49	MOTIFS
					ArgE / dapE / ACY1 / CPG2 / yscS family signature 2: A76-G115	MOTIFS
53	7509968CD1	180	S23 T156	N30	signal_cleavage: M1-S20	SPSCAN
					Signal Peptide: M1-S16, M1-G18, M1-S20, M1-Q21, M1-E22, M1-S24	HMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Trypsin-like serine protease: F35-P180	HMMER_SMART
					Trypsin: L36-T165	HMMER_PFAM
					Serine proteases, trypsin family, histidine proteins BL00134: C61-C77	BLIMPS_BLOCKS
					Serine proteases, trypsin family, active sites: L53-G97	PROFILESCAN
					Chymotrypsin serine protease family (S1) signature PR00722: G62-C77, N120-V134	BLIMPS_PRINTS
					PROTEASE SERINE PRECURSOR SIGNAL HYDROLASE ZYMOGEN GLYCOPROTEIN FAMILY MULTIGENE FACTOR PD000046: G38-T137	BLAST_PRODOM
					TRYPSIN DM00018|P20151|25-256: G38-P173 DM00018|P07288|25-256: G38-R175 DM00018|P33619|25-256: G38-R175 DM00018|P12323|1-234: G38-P173	BLAST_DOMO
					Serine proteases, trypsin family, histidine active site: L72-C77	MOTIFS
54	7510232CD1	293	S21 S109 S145 S151 S214 S228 S273 T81 T224 T280	N58 N200 N278	Death effector domain: M1-N80, L131-Y210	HMMER_SMART
					Death effector domain: D2-M86, L132-K215	HMMER_PFAM
					ICE-like protease (caspase) p20 domain: P249-A284	HMMER_PFAM
					Caspase family histidine proteins BL01121: G251-F261	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Death effector domain PF01335: Q13-L65, R68-E85	BLIMPS_PFAM
					APOPTOTIC CASPASE 8 PRECURSOR ICE-LIKE PROTEASE MORT1-ASSOCIATED CED3 HOMOLOG MACH FADD HOMOLOGOUS PD036164: F213-P249	BLAST_PRODOM
					PROTEIN APOPTOSIS CASH APOPTOTIC HOMOLOG CASPASE IFL ICE ISOFORM DEATH CASPASE 8 PD004333: L7-E84	BLAST_PRODOM
55	7510233CD1	511	S21 S109 S145 S151 S219 S224 S229 S243 S288 S337 S348 S443 S460 S483 T81 T239 T295 T304 T305 T405 T501 Y453	N58 N200 N293	Caspase, interleukin-1 beta converting enzyme: V257-S510	HMMER_SMART
					Death effector domain: M1-N80, L131-Y210	HMMER_SMART
					Death effector domain: D2-M86, L132-K215	HMMER_PFAM
					ICE-like protease (caspase) p10 domain: Q421-P509	HMMER_PFAM
					ICE-like protease (caspase) p20 domain: P264-K399	HMMER_PFAM
					Caspase family histidine proteins BL01121: G266-F276, I289-I324, D335-G350, C377-G394, Y424-C458, L475-D487	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Interleukin-1B converting enzyme signature PR00376: P264-A277, F342-G350, C377-D395, S443-I454	BLIMPS_PRINTS
					Death effector domain PF01335: Q13-L65, R68-E85	BLIMPS_PFAM
					PRECURSOR PROTEASE HYDROLASE THIOL ZYMOGEN APOPTOSIS PROTEIN APOPTOTIC CASPASE 1 CYSTEINE PD001408: R265-D395	BLAST_PRODOM
					APOPTOTIC CASPASE 8 PRECURSOR ICE-LIKE PROTEASE MORT1-ASSOCIATED CED3 HOMOLOG MACH FADD HOMOLOGOUS PD036164: F228-P264	BLAST_PRODOM
					PRECURSOR PROTEASE HYDROLASE THIOL ZYMOGEN APOPTOSIS PROTEIN APOPTOTIC CASPASE 1 CYSTEINE PD007531: I425-F508	BLAST_PRODOM
					PROTEIN APOPTOSIS CASH APOPTOTIC HOMOLOG CASPASE IFL ICE ISOFORM DEATH CASPASE 8 PD004333: L7-E84	BLAST_PRODOM
					INTERLEUKIN-1 BETA CONVERTING ENZYME FAMILY HISTIDINE DM01067|P42574|7-177: T253-S407 DM01067|P55210|34-204: Y258-T405 DM01067|P55211|128-311: D255-E409 DM01067|P55212|8-184: Y258-V403	BLAST_DOMO
					Caspase family cysteine active site: K383-G394	MOTIFS
					Caspase family histidine active site: H336-G350	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
56	7510304CD1	429	S90 S101 S114 S148 S191 T84 T112 T140 T378 T418	N124	signal_cleavage: M1-A38	SPSCAN
					Signal Peptide: M1-A38, M1-A34	HMMER
					Trypsin-like serine protease: R174-I412	HMMER_SMART
					Trypsin: I175-I412	HMMER_PFAM
					Cytosolic domain: M1-A19 Transmembrane domain: A20-I39 Non-cytosolic domain: G40-L429	TMHMMER
					Kringle domain proteins BL00021: C200-F217, V288-G309, C371-I412	BLIMPS_BLOCKS
					Serine proteases, trypsin family, histidine proteins BL00134: C200-C216, D359-R382, P399-I412	BLIMPS_BLOCKS
					Type I fibronectin domain proteins BL01253: C200-A213, S277-Y313, Y314-G352, I358-C371, W381-A415	BLIMPS_BLOCKS
					Serine proteases, trypsin family, active sites: L192-G241	PROFILESCAN
					Serine proteases, trypsin family, active sites: I344-A394	PROFILESCAN
					Chymotrypsin serine protease family (S1) signature PR00722: G201-C216, E265-L279, I358-V370	BLIMPS_PRINTS
					REPEAT PRECURSOR GLYCOPROTEIN PD00120: G201-A213, D269-V273, D359-G367	BLIMPS_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					PROTEASE SERINE HEPSIN TRANSMEMBRANE HYDROLASE SIGNAL ANCHOR PD021735: K4-R174	BLAST_PRODOM
					PROTEASE SERINE PRECURSOR SIGNAL HYDROLASE ZYMOGEN GLYCOPROTEIN FAMILY MULTIGENE FACTOR PD000046: I175-P399, L244-I412	BLAST_PRODOM
					TRYPSIN DM00018|P05981|163-403: I175-I416 DM00018|P98072|800-1033: R174-I412 DM00018|P14272|391-624: I175-I416 DM00018|P26262|391-624: I175-Q414	BLAST_DOMO
					Serine proteases, trypsin family, histidine active site: L211-C216	MOTIFS
					Serine proteases, trypsin family, serine active site: D359-V370	MOTIFS
57	7510461CD1	412	S11 S141 S270 S332 S412 T135 T196 T322 T341 T396	N173 N320 N339	signal_cleavage: M1-R33	SPSCAN
					Cytosol aminopeptidase family, catalytic domain: V178-S412	HMMER_PFAM
					Cytosol aminopeptidase proteins BL00631: D186-E199, I255-A288, A313-V353, I365-W380	BLIMPS_BLOCKS
					Cytosol aminopeptidase signature PR00481: I255-K272, M277-I298, E314-T335, V336-A356, I365-W380	BLIMPS_PRINTS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					AMINOPEPTIDASE HYDROLASE LEUCINE ZINC LAP LEUCYL CYTOSOL PROLINE PROLYL PROTEIN PD002804: F121-A378	BLAST_PRODOM
					CYTOSOL AMINOPEPTIDASE DM01145|P34629|152-482: L171-Q375 DM01145|P47631|117-443: V178-G373 DM01145|Q10401|1-338: V178-A378 DM01145|P47707|130-457: D176-A378	BLAST_DOMO
					ATP/GTP-binding site motif A (P-loop): G268-T275	MOTIFS
					Cytosol aminopeptidase signature: N340-L347	MOTIFS
58	7510392CD1	123	S62 T61 T78 T99		signal_cleavage: M1-A18	SPSCAN
					Signal Peptide: M1-A18, M1-S20, M1-G22, M1-C26	HMMER
					Domain in ryanodine and inositol trisphosphate receptors: L21-K75, G83-L123	HMMER_SMART
					MIR domain: L21-K75, G83-L123	HMMER_PFAM
					Cytosolic domain: L33-L123 Transmembrane domain: G10-L32 Non-cytosolic domain: M1-L9	TMHMMER

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
59/7313196CB1/ 1389	1-32, 1-34, 1-267, 31-336, 31-345, 31-416, 31-517, 31-542, 31-626, 31-631, 31-636, 31-638, 31-674, 31-691, 31-706, 31-864, 31-878, 41-659, 51-289, 93-647, 250-283, 636-1251, 708-1389, 734-1170
60/6465289CB1/ 4131	1-311, 1-570, 18-207, 25-803, 50-264, 50-594, 263-821, 381-916, 501-978, 501-1236, 501-1247, 511-1165, 519-1388, 526-739, 980-1718, 1064-1269, 1064-1449, 1067-1408, 1174-1445, 1344-1966, 1407-1822, 1522-1688, 1522-1797, 1549-1901, 1602-2081, 1660-1926, 1660-2422, 1664-2425, 1689-2384, 1715-1992, 1789-2056, 1789-2262, 1789-2349, 1832-2383, 1861-2458, 1928-2197, 1929-2190, 2001-2161, 2012-2576, 2024-2755, 2035-2602, 2053-2632, 2078-2805, 2099-2875, 2127-2901, 2161-2746, 2172-2782, 2214-2502, 2231-2731, 2239-2832, 2279-2676, 2287-2860, 2293-2979, 2295-2908, 2298-2833, 2320-2864, 2358-2621, 2371-2903, 2371-2947, 2388-2624, 2390-2884, 2392-2947, 2409-2699, 2420-2932, 2420-2958, 2426-3041, 2450-2789, 2476-3184, 2477-3144, 2478-2782, 2508-2779, 2550-3172, 2595-3101, 2613-3119, 2618-2815, 2652-3195, 2680-3681, 2707-3260, 2757-3009, 2759-3413, 2761-3333, 2810-3068, 2854-3124, 2901-3165, 2901-3453, 2914-3188, 2982-3290, 3004-3344, 3045-3255, 3045-3623, 3251-3530, 3278-3326, 3317-3545, 3321-3558, 3321-3808, 3355-3673, 3373-3610, 3388-3698, 3391-3628, 3417-3649, 3417-3696, 3460-3733, 3488-3641, 3503-3773, 3504-3765, 3515-3835, 3518-3830, 3530-3754, 3546-4131, 3652-3843, 3682-3942, 3682-3960, 3780-4028
61/7506357CB1/ 3249	1-231, 1-235, 22-260, 22-261, 22-263, 22-265, 22-269, 22-271, 22-276, 22-278, 22-284, 22-289, 22-290, 22-299, 22-515, 22-610, 22-632, 22-660, 22-676, 22-761, 22-3237, 24-328, 26-275, 30-272, 31-678, 32-262, 41-686, 57-270, 60-267, 60-295, 63-689, 71-592, 73-495, 77-729, 83-359, 83-600, 99-326, 107-755, 110-676, 111-753, 128-403, 140-692, 141-735, 142-271, 175-729, 221-796, 224-756, 234-847, 289-777, 298-872, 332-881, 338-845, 382-983, 412-893, 418-917, 477-1360, 507-1075, 515-1125, 551-969, 551-1048, 551-1091, 554-1060, 571-880, 624-735, 642-756, 711-1256, 755-874, 755-1168, 755-1181, 766-1176, 770-1000, 779-1126, 784-1373, 787-1296, 799-1083, 825-1254, 900-1022, 908-1172, 922-1287, 963-1288, 977-1596, 990-1356, 990-1516, 990-1578, 990-1593, 990-1601, 990-1625, 990-1664, 990-1747, 990-1778, 999-1265, 999-1615, 1025-1176, 1028-1281, 1050-1690, 1087-1181, 1087-1720, 1088-1361, 1136-1899, 1138-1768, 1281-2009, 1284-1604, 1285-1590, 1308-1803, 1331-1460, 1331-1467, 1346-2056, 1370-2058, 1375-2076,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1380-1656, 1391-2056, 1401-2035, 1416-1959, 1473-2203, 1476-2104, 1558-2057, 1565-1951, 1571-1811, 1571-2016, 1589-2119, 1606-1891, 1616-1835, 1619-2255, 1625-2184, 1687-2057, 1721-1977, 1736-2115, 1745-2349, 1772-2035, 1815-2253, 1840-2087, 1869-2145, 1888-2049, 1890-2181, 1896-2159, 1932-2134, 1984-2177, 1998-2202, 2018-2223, 2044-2253, 2152-2442, 2179-2459, 2251-2497, 2253-2472, 2253-2806, 2253-2830, 2255-2562, 2255-2740, 2256-2469, 2256-2768, 2260-2381, 2260-2511, 2260-2516, 2260-2527, 2261-2930, 2262-2498, 2265-2667, 2270-2800, 2276-2458, 2276-2506, 2287-2514, 2295-2533, 2297-2530, 2310-2785, 2311-2699, 2311-2968, 2316-2574, 2316-2609, 2318-2803, 2327-2590, 2337-2793, 2337-2843, 2339-3053, 2375-2996, 2376-2557, 2376-2809, 2384-2593, 2386-2788, 2386-2903, 2388-2648, 2393-2684, 2398-2760, 2400-3227, 2401-2645, 2412-2686, 2412-2917, 2416-2964, 2418-3092, 2426-3137, 2438-2916, 2438-3127, 2442-2688, 2467-2601, 2474-3176, 2477-3066, 2481-2940, 2485-2740, 2485-2743, 2485-3176, 2499-2684, 2503-3037, 2507-3154, 2507-3183, 2507-3190, 2516-3178, 2522-2754, 2528-2766, 2529-2776,
	2531-2765, 2541-3084, 2546-2738, 2550-2850, 2559-2870, 2563-3156, 2566-3115, 2577-2804, 2578-2841, 2579-3240, 2585-2868, 2585-2869, 2586-3125, 2594-2894, 2595-2975, 2595-2997, 2597-2808, 2597-2853, 2610-2843, 2617-3052, 2633-2893, 2636-3185, 2636-3238, 2637-2891, 2641-2944, 2643-2882, 2643-3184, 2653-3215, 2656-2825, 2657-2936, 2665-3190, 2675-3099, 2688-3248, 2702-3240, 2709-2964, 2715-2949, 2715-3039, 2719-2972, 2732-3238, 2751-3089, 2752-3225, 2754-3228, 2755-3029, 2759-3063, 2763-3225, 2772-3243, 2774-3227, 2776-3221, 2778-3001, 2778-3034, 2778-3227, 2779-3026, 2780-3220, 2781-3243, 2787-2904, 2787-3039, 2787-3045, 2799-3081, 2806-3243, 2807-3219, 2811-3041, 2813-3225, 2814-3227, 2815-3220, 2815-3227, 2820-3225, 2823-3225, 2825-3229, 2827-3225, 2828-3239, 2829-3193, 2829-3225, 2830-3193, 2831-3219, 2852-3054, 2866-3224, 2870-3238, 2872-3077, 2873-3227, 2882-3150, 2887-3194, 2887-3240, 2888-3207, 2901-3225, 2910-3094, 2922-3225, 2927-3249, 2964-3220, 2968-3205, 2984-3225, 2986-3225, 2986-3239, 3000-3232, 3006-3193, 3011-3239, 3066-3213, 3066-3249
62/6878857CB1/ 986	1-656, 40-746, 186-706, 203-731, 374-986, 380-668, 381-781, 383-779, 394-677, 395-775, 396-776, 397-774, 403-775, 404-772, 408-780, 420-694, 422-699, 440-773, 441-775, 444-780, 450-780, 455-775, 456-787, 458-772, 461-776

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
63/7506021CB1/ 3665	1-600, 2-506, 31-506, 114-506, 132-341, 523-1087, 523-3634, 531-1004, 615-1116, 778-991, 1204-1465, 1204-1752, 1374-2042, 1582-1799, 1583-2214, 1630-2257, 1663-2095, 1671-2454, 1671-2463, 1672-2460, 1716-2460, 1874-2460, 1936-2241, 1936-2452, 1936-2455, 1936-2458, 1936-2459, 1936-2473, 1936-2507, 1975-2707, 1977-2460, 1980-2505, 2138-2416, 2138-2642, 2194-2499, 2230-2567, 2230-2605, 2250-2693, 2335-2639, 2363-2719, 2403-2645, 2584-3069, 2718-3069, 2718-3344, 2719-3180, 2720-3161, 2726-3180, 2736-3343, 2761-3188, 2771-3188, 2771-3257, 2777-3187, 2777-3339, 2792-3344, 2801-3344, 2803-3180, 2804-3344, 2805-3344, 2815-3317, 2824-3348, 2826-3339, 2835-3330, 2856-3308, 2882-3146, 2890-3344, 2891-3344, 2908-3344, 2925-3344, 2940-3343, 2958-3344, 3032-3209, 3045-3150, 3086-3310, 3121-3590, 3126-3548, 3175-3665, 3220-3658, 3221-3454, 3236-3513, 3304-3634, 3401-3633, 3477-3620
64/7503356CB1/ 1440	1-336, 1-403, 1-455, 1-468, 1-519, 1-1440, 25-263, 25-414, 25-443, 34-569, 34-907, 39-722, 62-779, 528-1165, 552-1163, 567-1171, 571-1165, 576-1171, 591-1165, 609-1024, 635-1165, 663-1165, 663-1171, 669-1165, 674-1165, 704-1096, 714-1165, 763-1165, 796-1267, 848-1147, 940-1096, 959-1096, 1096-1417
65/7509052CB1/ 1064	1-210, 1-224, 1-908, 3-255, 3-269, 4-251, 5-279, 9-591, 9-629, 9-631, 15-465, 20-454, 30-580, 46-399, 74-595, 77-728, 99-850, 222-757, 263-700, 265-524, 265-537, 267-898, 268-912, 269-863, 271-558, 272-508, 278-918, 281-773, 283-538, 285-507, 285-524, 285-557, 285-571, 285-864, 288-912, 291-913, 292-827, 293-523, 295-827, 295-920, 299-767, 301-862, 310-806, 311-913, 316-510, 321-948, 323-565, 325-895, 326-616, 326-630, 327-612, 328-844, 329-602, 336-632, 339-897, 339-913, 344-915, 349-598, 355-636, 355-836, 358-654, 359-727, 366-556, 368-897, 369-908, 372-895, 372-896, 376-628, 376-909, 378-656, 378-908, 381-739, 381-770, 382-647, 383-913, 384-656, 386-628, 388-629, 388-652, 388-667, 389-861, 390-653, 391-676, 392-863, 396-913, 398-682, 407-916, 409-914, 410-912, 411-912, 412-894, 414-678, 414-701, 415-894, 416-706, 417-765, 417-890, 418-675, 418-681, 418-717, 419-630, 419-695, 419-741, 419-765, 420-653, 420-752, 421-704, 421-734, 421-775, 423-887, 423-895, 424-888, 425-701, 427-896,
	429-895, 429-913, 430-897, 434-723, 434-724, 434-911, 435-701, 436-628, 436-685, 436-689, 436-695, 436-905, 439-678, 439-873, 442-681, 443-912, 444-912, 445-734, 445-884, 446-897, 447-528, 448-684, 449-922, 450-912, 453-913, 454-844, 455-917, 456-888, 459-895, 459-908, 461-908, 463-883, 464-879, 464-912, 469-912, 470-700, 474-895, 475-761, 476-895, 476-913, 477-894, 478-896, 479-722, 479-770, 479-908, 479-914, 479-920, 480-858, 480-897, 481-896, 481-897, 482-895, 483-895, 483-896, 483-897, 484-672, 485-892, 485-897, 485-916, 486-888, 487-770, 487-888, 488-693, 488-706, 488-896, 488-897, 489-830, 489-878, 489-895, 490-892, 491-888, 491-897, 491-911, 492-729, 492-895, 492-898, 493-753, 493-795, 493-898, 493-910, 494-877, 494-895, 495-897, 495-914, 498-804, 499-887, 500-755, 500-789, 500-895, 502-988, 504-777, 504-897, 505-897, 505-922, 507-851, 508-752, 508-896, 509-854, 510-893, 510-897, 511-891, 511-895, 511-897, 512-909, 513-715, 514-836, 514-896, 514-913, 515-897, 518-859, 518-895, 519-731,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	519-895, 523-897, 526-757, 526-768, 526-895, 527-895, 530-897, 532-778, 533-870, 533-897, 536-793, 536-804, 536-897, 542-829, 542-887, 543-804, 543-836, 546-793, 546-848, 546-920, 547-789, 547-895, 547-896, 549-658, 549-781, 549-895, 549-909, 549-913, 550-990, 551-869, 551-909, 553-796, 555-897, 555-909, 556-801, 558-909, 558-935, 559-805, 561-799, 561-897, 562-804, 562-897, 563-828, 563-838, 564-787, 565-803, 565-807, 567-759, 567-895, 569-895, 569-910, 569-913, 571-897, 572-847, 572-897, 573-895, 573-897, 577-794, 577-826, 577-853, 577-892, 577-896, 578-848, 581-807, 581-812, 581-864, 582-895, 583-982, 584-861, 584-888, 585-882, 587-837, 587-861, 588-859, 588-895, 589-830, 589-834, 590-897, 592-888, 593-873, 594-847, 597-813, 597-825, 598-828, 598-896, 600-874, 600-895, 608-822, 608-830, 610-895, 612-880, 612-887, 613-718, 614-834, 615-895, 615-897, 616-897, 619-845, 619-879, 619-892, 621-873, 621-888, 621-897, 623-877, 624-889, 624-896, 625-877, 629-887, 629-888, 629-913, 629-914, 630-896, 632-987,
	633-897, 633-908, 633-925, 636-1015, 637-897, 641-898, 641-909, 642-858, 646-894, 651-898, 652-893, 652-920, 653-898, 653-905, 655-910, 656-908, 658-766, 659-891, 659-895, 659-896, 659-897, 659-898, 659-904, 659-908, 662-908, 665-895, 667-892, 668-905, 669-895, 669-940, 669-992, 675-911, 677-909, 681-940, 683-898, 683-909, 684-910, 688-895, 688-918, 690-910, 693-895, 697-894, 700-895, 701-908, 709-898, 710-846, 710-909, 717-945, 719-910, 719-911, 719-975, 721-940, 721-987, 725-895, 728-897, 728-910, 729-892, 729-898, 729-906, 729-931, 729-932, 733-895, 734-910, 738-909, 738-998, 740-910, 741-906, 757-908, 767-943, 770-897, 771-1013, 787-911, 802-893, 818-896, 820-945, 820-1064, 848-887
66/7503366CB1/ 4497	1-240, 1-357, 1-392, 1-412, 1-425, 1-430, 1-467, 1-484, 1-522, 1-545, 1-560, 1-565, 1-567, 1-571, 1-575, 1-577, 1-586, 1-588, 1-594, 1-600, 1-623, 1-627, 1-633, 1-635, 1-638, 1-642, 1-644, 1-649, 1-654, 1-657, 1-667, 1-671, 1-673, 1-675, 1-677, 1-730, 1-753, 1-4497, 27-292, 27-571, 28-275, 28-654, 31-705, 39-788, 71-596, 73-596, 119-788, 163-689, 227-847, 239-1110, 256-826, 262-828, 262-844, 285-497, 285-498, 288-497, 295-488, 401-1219, 520-1342, 538-1226, 551-1226, 578-1342, 611-1342, 616-1166, 617-1342, 619-1341, 623-1342, 627-1342, 629-1342, 691-1336, 699-1336, 723-1342, 732-1339, 733-1421, 746-1421, 750-1289, 795-1490, 798-1477, 805-1342, 832-1342, 914-1618, 948-1618, 956-1613, 975-1561, 1064-1683, 1077-1618, 1232-1487, 1232-1763, 1232-1787, 1232-1802, 1232-1862, 1302-1585, 1462-1993, 1543-2192, 2332-2914, 2713-3186, 2877-3350, 3165-3300, 3384-3987, 3389-3872, 3430-3809, 3447-4038, 3453-3862, 3463-3664, 3463-3682, 3463-3885, 3494-3752, 3496-4284, 3535-4077, 3553-4153, 3577-3807,
	3578-4133, 3598-4102, 3614-3853, 3614-4284, 3622-3941, 3625-4235, 3637-3875, 3639-3846, 3639-4141, 3661-4278, 3685-3964, 3692-3946, 3692-4279, 3712-3992, 3744-4279, 3755-4007, 3755-4235, 3756-4301, 3795-4090, 3798-3941, 3798-4021, 3805-4046, 3819-4118, 3819-4497, 3835-4306, 3863-4128, 3865-4299, 3874-4147, 3941-4308, 3973-4302, 3981-4301, 4023-4304, 4086-4279, 4086-4283, 4094-4309, 4205-4278, 4216-4287, 4216-4290

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
67/7505933CB1/ 1247	1-226, 1-243, 1-418, 1-1231, 10-314, 10-330, 11-275, 12-279, 12-288, 12-309, 16-223, 16-224, 16-248, 16-269, 16-272, 16-279, 16-289, 16-294, 17-247, 17-297, 17-346, 18-250, 20-231, 21-342, 22-237, 22-254, 23-144, 23-230, 23-255, 23-269, 24-218, 24-267, 24-268, 24-276, 24-284, 24-286, 24-297, 24-301, 24-316, 24-346, 26-255, 26-259, 26-268, 26-316, 26-320, 26-329, 26-330, 27-176, 28-426, 29-416, 30-271, 30-278, 30-290, 30-326, 30-327, 31-287, 31-306, 32-277, 33-236, 33-242, 33-273, 33-274, 33-323, 34-224, 34-245, 34-256, 34-290, 34-292, 34-313, 34-320, 34-321, 34-357, 34-360, 35-229, 35-230, 35-262, 35-267, 35-268, 35-275, 35-288, 35-290, 35-300, 35-303, 35-315, 35-319, 35-345, 36-162, 36-284, 36-328, 37-129, 37-147, 37-185, 37-203, 37-204, 37-263, 37-264, 37-276, 37-279, 37-281, 37-287, 37-298, 37-309, 37-327, 37-333, 37-334, 37-337, 37-350, 37-426, 38-196, 38-277, 38-278, 38-283, 38-294, 38-303, 38-317, 39-229, 39-269, 39-300, 39-385, 40-290, 40-303, 40-330, 41-253, 41-271, 41-306,
	41-320, 41-324, 41-341, 42-219, 42-272, 42-293, 42-300, 42-319, 42-334, 42-343, 42-348, 42-355, 42-367, 42-370, 43-201, 43-242, 43-363, 43-365, 44-261, 44-290, 44-293, 44-295, 44-306, 44-309, 44-323, 44-337, 44-346, 44-390, 45-288, 45-292, 45-300, 45-309, 45-319, 45-327, 45-353, 45-359, 45-416, 46-164, 46-251, 46-321, 46-427, 47-294, 48-293, 48-298, 48-299, 48-311, 48-312, 48-314, 48-339, 48-426, 50-265, 51-350, 52-283, 52-286, 52-294, 52-321, 52-341, 52-376, 53-283, 55-262, 55-271, 55-302, 55-307, 55-310, 55-326, 55-343, 55-349, 58-220, 58-236, 60-160, 60-305, 61-328, 62-299, 63-328, 63-332, 68-345, 68-380, 75-220, 75-390, 77-323, 77-341, 82-367, 82-426, 87-367, 96-393, 97-354, 97-373, 97-383, 105-215, 105-373, 110-348, 121-415, 128-352, 130-418, 157-732, 163-426, 172-426, 174-401, 174-426, 177-381, 177-426, 178-413, 180-357, 180-370, 184-426, 191-333, 194-378, 300-426, 376-606, 414-657, 419-582, 419-620, 419-635, 419-642, 419-665, 419-685, 419-691, 419-712, 419-715, 419-728, 419-946, 419-1189,
	421-556, 421-585, 421-675, 421-714, 422-558, 422-762, 422-1158, 429-708, 429-865, 430-915, 430-1031, 435-1139, 436-974, 439-721, 443-623, 443-701, 443-721, 447-1086, 448-863, 449-700, 452-1082, 452-1125, 453-691, 456-1100, 457-1008, 458-686, 464-696, 464-726, 465-652, 466-697, 466-711, 466-778, 467-1088, 468-1182, 471-709, 473-732, 473-1024, 473-1054, 476-731, 476-780, 476-908, 477-744, 481-1161, 485-1113, 486-784, 486-795, 488-745, 488-751, 493-1061, 497-849, 499-723, 503-842, 504-1174, 506-1161, 513-654, 525-666, 530-785, 534-930, 539-787, 539-800, 539-827, 539-906, 549-960, 551-714, 551-820, 551-857, 552-805, 552-817, 552-857, 557-1130, 567-844, 568-810, 572-1247, 579-819, 579-827, 585-805, 585-859, 585-868, 585-1203, 586-1143, 590-809, 591-1122, 591-1175, 591-1176, 593-915, 595-1212, 595-1230, 597-1229, 597-1230, 599-840, 602-875, 602-883, 602-886, 602-1218, 605-822, 609-852, 609-1231, 615-868, 615-1230, 616-1216, 620-819, 626-910, 626-987, 630-1188, 631-885,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	631-887, 633-1154, 634-935, 640-885, 642-914, 642-925, 644-1231, 650-1247, 661-842, 661-898, 661-904, 661-911, 661-916, 661-921, 661-1230, 661-1247, 662-1222, 662-1231, 662-1247, 663-1070, 664-1247, 667-951, 667-956, 671-1247, 672-1213, 673-1095, 676-930, 676-1233, 679-1146, 679-1245, 685-1209, 686-1245, 688-942, 702-1231, 703-1247, 709-1230, 711-1233, 715-958, 725-1004, 727-1231, 732-1003, 738-1063, 739-908, 743-1236, 746-1031, 747-1015, 749-1246, 756-999, 765-1052, 766-975, 773-1054, 775-1231, 777-1049, 779-1231, 784-1231, 785-1129, 785-1237, 786-1229, 787-1231, 788-1231, 789-1231, 791-1231, 791-1247, 794-1247, 795-1247, 798-1039, 800-1173, 800-1229, 805-1230, 806-1247, 809-1074, 811-1235, 811-1247, 814-1232, 815-1231, 819-1208, 820-1228, 821-1230, 821-1231, 823-1108, 823-1227, 824-1231, 825-1231, 825-1247, 826-1052, 827-1089, 827-1218, 827-1246, 828-1231, 829-1231, 833-1247, 836-1231, 839-1231, 841-1066, 842-1100, 843-1231, 843-1235, 847-1231, 851-1231, 852-1054, 853-1230, 855-1175, 855-1177, 857-1231, 857-1235, 859-1053, 865-1136, 865-1231, 868-1146, 876-1231, 878-1231,
	879-947, 879-1231, 880-1230, 886-1121, 892-1162, 892-1166, 892-1185, 906-1231, 908-1024, 908-1165, 909-1231, 923-1231, 929-1231, 931-1181, 935-1235, 936-1209, 939-1231, 940-1235, 943-1231, 944-1229, 945-1231, 945-1233, 946-1209, 946-1214, 946-1216, 948-1229, 959-1211, 962-1231, 968-1247, 972-1160, 973-1188, 973-1231, 973-1247, 974-1169, 974-1247, 975-1229, 975-1233, 982-1231, 992-1115, 992-1208, 992-1247, 1016-1231, 1018-1236, 1022-1247, 1024-1239, 1030-1247, 1049-1247, 1058-1231, 1059-1231, 1060-1231, 1062-1246, 1081-1231, 1084-1247, 1092-1222, 1092-1227, 1092-1230, 1092-1235, 1099-1235, 1116-1231, 1122-1231, 1126-1247, 1130-1247, 1138-1227
68/7507064CB1/ 714	1-683, 1-707, 92-529, 125-631, 143-590, 143-707, 167-714, 229-704, 267-705, 304-714, 358-712, 373-707, 552-714, 553-714, 562-714
69/1439986CB1/ 1008	1-164, 49-236, 49-429, 49-439, 49-527, 49-530, 49-542, 49-660, 49-666, 49-708, 49-716, 49-806, 106-505, 109-518, 134-1008, 158-1003, 264-936, 267-608, 268-918

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
70/2008979CB1/ 2425	1-596, 1-2425, 129-379, 132-425, 132-621, 267-797, 308-820, 310-438, 311-892, 396-868, 431-1060, 439-919, 441-860, 448-841, 510-809, 518-1138, 524-796, 606-1079, 607-807, 614-824, 632-1164, 637-889, 690-1224, 708-920, 708-952, 708-959, 708-1157, 719-1204, 773-1030, 788-1203, 811-1323, 838-1201, 840-1203, 887-1392, 896-1429, 938-1203, 949-1224, 959-1213, 972-1224, 972-1233, 973-1165, 974-1178, 992-1209, 1012-1197, 1024-1266, 1024-1474, 1028-1321, 1028-1607, 1030-1281, 1037-1203, 1130-1412, 1133-1429, 1137-1419, 1158-1412, 1163-1407, 1176-1437, 1204-1453, 1209-1516, 1215-1475, 1227-1780, 1265-1545, 1267-1575, 1274-1779, 1274-1872, 1289-1590, 1292-1692, 1297-1552, 1308-1624, 1320-1780, 1326-1568, 1338-1570, 1338-1773, 1363-1624, 1394-1635, 1394-1676, 1395-1611, 1396-1676, 1412-1672, 1432-1909, 1441-1698, 1464-1730, 1500-1716, 1563-2139, 1630-2102, 1658-1855, 1697-2140, 1728-2005, 1759-2122, 1759-2132, 1766-1995, 1776-2137, 1832-2103, 1832-2345, 1844-2124, 1984-2215, 1984-2236, 2084-2422, 2113-2351, 2113-2371, 2187-2425, 2206-2425
71/90073157CB1/ 856	1-776, 4-616, 30-856, 640-856
72/7506782CB1/ 1318	1-256, 7-293, 7-728, 7-1296, 76-368, 77-376, 129-423, 176-474, 207-484, 207-497, 210-502, 250-503, 278-496, 505-653, 505-739, 505-1112, 505-1132, 505-1286, 520-1044, 528-1112, 536-1112, 536-1286, 543-771, 546-1154, 557-1277, 558-1086, 572-775, 572-1254, 573-809, 574-1246, 577-838, 580-787, 589-884, 589-1135, 596-832, 607-998, 611-1098, 611-1257, 632-730, 635-1104, 638-873, 648-1188, 649-921, 649-1214, 665-853, 682-968, 691-881, 693-1194, 711-1255, 722-988, 723-985, 727-1298, 733-980, 737-1308, 757-1313, 799-1296, 815-1062, 819-1287, 832-1098, 847-1292, 848-1315, 849-1298, 856-1318, 863-1116, 871-1296, 871-1297, 873-1316, 874-1296, 879-1296, 893-1298, 896-1298, 901-1298, 909-1296, 918-1171, 918-1311, 918-1312, 918-1313, 919-1312, 932-1296, 966-1298, 967-1296, 979-1294, 979-1298, 982-1188, 984-1224, 984-1251, 984-1308, 988-1295, 989-1251, 1006-1318, 1042-1298, 1043-1180, 1043-1256, 1043-1298, 1046-1298, 1049-1318, 1089-1318, 1144-1318, 1194-1318
73/7506941CB1/ 2251	1-495, 1-2249, 2-499, 39-301, 74-780, 74-792, 74-803, 74-841, 74-851, 74-862, 74-898, 75-898, 78-844, 85-869, 147-754, 274-644, 343-603, 433-683, 433-897, 460-745, 572-901, 599-1454, 898-1211, 898-1236, 898-1265, 898-1314, 898-1707, 901-1616, 919-1662, 939-1192, 939-1462, 940-1134, 940-1210, 940-1224, 940-1544, 940-1545, 940-1581, 958-1416, 968-1552, 993-1279, 1027-1294, 1110-1381, 1110-1402, 1133-1395, 1134-2153, 1138-1701, 1148-1373, 1150-1377, 1171-2153, 1173-1677, 1188-1764, 1188-1873, 1198-1425, 1198-1720, 1204-1702, 1222-1796, 1235-2153, 1262-1752, 1297-2153, 1299-1438, 1340-2153, 1341-2153, 1344-2153, 1360-1924, 1374-1862, 1403-2153, 1404-2153, 1458-2153, 1476-1762, 1496-2152, 1504-1638, 1511-1717, 1686-2232, 1719-2137, 1738-1875, 1808-2251, 1877-2118
74/7507072CB1/ 1147	1-482, 1-709, 1-739, 1-1140, 210-1069, 425-1069, 461-1069, 533-1069, 626-880, 626-1032, 626-1133, 770-1140, 771-1139, 773-1140, 832-1140, 835-1098, 855-1140, 957-1147, 966-1147, 967-1147, 969-1146, 969-1147, 1018-1099

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
75/7507083CB1/ 800	1-414, 1-482, 1-510, 1-532, 1-535, 2-532, 5-535, 16-517, 18-96, 18-395, 18-441, 18-449, 18-535, 18-557, 18-651, 18-737, 18-785, 18-797, 18-798, 18-799, 18-800, 23-800, 25-567, 30-670, 63-535, 66-699, 75-707, 90-729, 103-725, 113-665, 114-776, 117-474, 123-709, 138-709, 148-800, 247-729, 493-606
76/7509097CB1/ 1296	1-820, 1-834, 20-588, 20-596, 20-606, 23-358, 23-434, 23-443, 23-454, 23-493, 23-579, 23-596, 23-599, 23-604, 23-610, 23-639, 23-833, 23-1296, 24-658, 24-697, 24-722, 27-429, 27-614, 41-385, 41-736, 41-780, 44-644, 44-705, 44-711, 44-832, 44-833, 44-882, 44-914, 47-805, 374-418, 392-596, 440-1296, 565-735, 1133-1296
77/7509118CB1/ 1407	1-820, 1-873, 23-1407, 41-780, 41-874, 44-834, 44-904, 44-923, 114-223, 114-874, 114-1059, 393-1015, 455-1126, 462-1126, 463-1126, 503-1124, 503-1126, 505-1126, 565-735, 568-1386, 579-1126, 603-861, 801-1126, 914-1125, 987-1185
78/7509312CB1/ 1448	1-859, 1-1448, 160-741, 229-434, 230-852, 241-959, 292-963, 299-963, 300-963, 311-936, 323-1095, 340-961, 340-963, 342-963, 367-1158, 402-572, 405-1158, 416-963, 440-698, 479-963, 488-1158, 525-963, 565-1207, 565-1427, 567-1357, 569-1427, 572-1427, 601-1427, 606-1427, 616-963, 618-1427, 629-1427, 638-963, 647-963, 647-1415, 647-1427, 651-1427, 654-1415, 654-1427, 670-1427, 727-1415, 751-962, 763-1415, 773-1050, 783-1427, 801-1427, 808-925, 809-1230, 809-1427, 810-1427, 824-1022, 982-1427, 1054-1448
79/90126902CB1/ 2360	1-615, 1-747, 4-762, 4-822, 98-343, 98-344, 212-502, 339-679, 368-1088, 470-1071, 485-936, 503-981, 550-1088, 568-1088, 608-1088, 618-1088, 695-1088, 741-1088, 769-1061, 769-1151, 769-1155, 769-1160, 769-1201, 769-1280, 769-1294, 985-1352, 995-1798, 1065-1376, 1128-1795, 1150-1798, 1195-1313, 1206-1368, 1208-1345, 1211-1798, 1321-1851, 1401-1757, 1401-1799, 1417-1768, 1417-1822, 1497-1798, 1567-2144, 1593-1671, 1593-1712, 1593-1933, 1640-2223, 1698-2216, 1767-1915, 1831-2335, 1880-2359, 1914-2239, 1914-2360, 1998-2360, 2082-2360, 2167-2360
80/7509352CB1/ 1109	1-283, 1-336, 1-816, 1-1104, 7-817, 330-1064, 330-1109, 361-1109, 363-1109, 375-1109, 378-1086, 384-1109, 385-1109, 387-1109, 411-643, 413-1109, 423-985, 448-1045, 487-1109, 498-1107, 503-1109, 525-1074, 540-933, 545-815, 571-1109, 578-825, 578-1056, 591-1109, 601-840, 627-888, 628-1109, 631-902, 631-1099, 633-918, 649-820, 653-1053, 725-1109, 735-1007, 751-1109, 767-1109, 786-1037, 816-964, 840-1097, 872-1108
81/7509341CB1/ 905	1-109, 1-539, 1-903, 7-458, 382-834, 539-834, 539-864, 540-898, 544-890, 547-890, 606-905, 651-864, 721-890, 800-898
82/7509367CB1/ 626	1-105, 1-301, 1-596, 6-301, 7-301, 17-301, 108-301, 185-301, 246-574, 302-626, 306-626, 309-626, 368-626, 413-626, 483-626, 562-626

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
83/7500455CB1/ 4095	1-4095, 12-4094, 72-577, 72-623, 72-653, 72-668, 72-674, 72-679, 72-705, 72-712, 72-723, 72-740, 72-758, 72-773, 72-782, 72-790, 72-802, 72-809, 72-812, 72-814, 96-775, 128-953, 138-901, 141-714, 172-823, 229-921, 280-574, 310-1080, 315-627, 354-927, 397-1027, 429-1086, 443-1035, 484-717, 484-1327, 486-607, 486-826, 502-1128, 542-1066, 582-1400, 658-1439, 806-1340, 813-1071, 813-1432, 876-1440, 890-1646, 902-1079, 922-1181, 922-1188, 929-1434, 1001-1467, 1037-1288, 1037-1340, 1037-1544, 1037-1603, 1041-1511, 1048-1531, 1079-1466, 1091-1233, 1114-1556, 1132-1555, 1163-1831, 1187-1467, 1191-1331, 1197-1468, 1248-1557, 1298-1506, 1301-1462, 1475-1774, 1475-1804, 1601-1907, 1623-2112, 1788-2191, 1804-2060, 1843-2308, 1906-2445, 1937-2301, 1955-2203, 1963-2553, 1963-2618, 2025-2250, 2025-2356, 2077-2566, 2146-2779, 2148-2406, 2207-2699, 2229-2495, 2247-2716, 2249-2729, 2276-2550, 2284-2493, 2290-2496, 2334-2965, 2339-2737, 2353-2601, 2353-3198, 2412-3030, 2416-2969, 2416-3025, 2433-3025, 2452-3025, 2453-2726, 2465-3025, 2489-3025, 2492-3025, 2498-3025, 2519-3011, 2534-2963, 2542-3025,
	2560-2887, 2568-2958, 2595-2859, 2601-3025, 2620-2866, 2620-2903, 2623-3025, 2696-3450, 2731-3473, 2741-3621, 2796-3052, 2799-3326, 2800-3315, 2818-3075, 2845-3628, 2849-3628, 2854-3624, 2857-3628, 2890-3628, 2908-3234, 2909-3626, 2932-3628, 3011-3330, 3017-3406, 3017-3598, 3017-3618, 3017-3644, 3020-3324, 3021-3290, 3022-3600, 3025-3254, 3025-3260, 3025-3264, 3025-3267, 3025-3270, 3025-3277, 3025-3280, 3025-3318, 3025-3319, 3025-3323, 3025-3324, 3025-3336, 3025-3391, 3025-3434, 3025-3448, 3025-3475, 3025-3503, 3025-3526, 3025-3529, 3025-3536, 3025-3553, 3025-3593, 3025-3603, 3025-3610, 3025-3616, 3025-3629, 3025-3642, 3025-3646, 3025-3648, 3025-3653, 3025-3663, 3025-3667, 3025-3677, 3025-3685, 3027-3638, 3034-3289, 3091-3502, 3099-3326, 3104-3870, 3105-3406, 3108-3327, 3111-3669, 3129-3427, 3137-3376, 3137-3385, 3137-3419, 3138-3406, 3150-3453, 3166-3434, 3175-3428, 3186-3451, 3191-3479, 3202-3456, 3227-4040, 3233-3842, 3243-3503, 3279-3515, 3307-3502, 3324-3549, 3326-3662, 3337-3565, 3382-3870, 3384-3614, 3386-3505, 3386-4086, 3388-3669, 3396-4084, 3411-3882, 3420-3730, 3435-3697, 3444-3676,
	3504-3737, 3504-3764, 3586-3874, 3590-4092, 3594-3812, 3618-4094, 3619-4094, 3631-4084, 3635-4087, 3640-4090, 3644-4092, 3647-4095, 3681-4088, 3695-4011, 3718-4001, 3723-4015, 3723-4088, 3723-4094, 3735-4011, 3776-3970, 3776-4089, 3826-4014, 3830-4067, 4010-4092, 4011-4084

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
84/7510401CB1/ 1308	1-125, 5-268, 17-230, 17-250, 17-259, 17-264, 17-280, 17-290, 17-315, 17-1308, 18-312, 19-269, 19-306, 20-309, 21-305, 21-340, 22-252, 24-312, 26-499, 28-267, 29-244, 29-258, 29-277, 29-299, 29-310, 29-346, 30-233, 31-274, 32-294, 32-306, 32-354, 38-218, 38-252, 38-256, 38-267, 38-274, 38-276, 38-279, 38-283, 38-284, 38-293, 38-295, 38-298, 38-311, 38-315, 38-326, 38-339, 40-166, 40-184, 40-290, 40-302, 41-234, 41-237, 41-249, 41-266, 41-275, 41-276, 41-277, 41-279, 41-281, 41-285, 41-286, 41-287, 41-295, 41-296, 41-312, 41-324, 41-334, 41-347, 42-273, 42-318, 42-326, 43-265, 43-290, 43-309, 43-335, 43-346, 44-300, 44-304, 44-308, 44-315, 44-317, 44-333, 45-336, 46-247, 46-293, 47-364, 51-305, 54-286, 56-265, 56-274, 56-278, 56-340, 58-289, 65-292, 89-265, 121-302, 121-311, 145-458, 162-694, 259-482, 278-428, 414-991, 469-667, 526-651, 569-961, 572-800, 573-821, 586-1209, 607-892, 646-874, 677-970, 719-979, 728-1304, 756-1308, 760-1308, 782-1052, 784-1308, 789-1225, 794-1213, 854-1090, 854-1304,
	863-1191, 866-1308, 870-1115, 870-1308, 871-1303, 873-1035, 873-1153, 875-1307, 876-1265, 876-1304, 876-1308, 878-1308, 879-1047, 881-1308, 887-1304, 887-1308, 890-1304, 892-1304, 894-1166, 900-1308, 906-1308, 910-1196, 913-1165, 922-1299, 927-1304, 929-1308, 930-1304, 930-1308, 935-1304, 962-1226, 963-1188, 965-1263, 976-1307, 977-1308, 993-1292, 995-1306, 996-1304, 1010-1308, 1013-1304, 1027-1308, 1046-1308, 1075-1308, 1076-1308, 1114-1301, 1162-1308, 1163-1300, 1173-1308, 1214-1304
85/7504702CB1/ 1196	1-138, 1-182, 1-186, 1-202, 1-222, 1-223, 1-225, 1-517, 1-519, 1-929, 1-1145, 2-220, 3-188, 4-397, 5-542, 5-577, 5-584, 7-206, 7-529, 10-231, 11-503, 11-585, 12-545, 22-542, 29-231, 30-540, 43-552, 51-346, 60-620, 60-745, 63-633, 66-578, 67-551, 69-638, 104-616, 108-358, 121-405, 125-666, 139-693, 162-731, 169-775, 184-722, 192-557, 231-511, 231-693, 231-849, 231-861, 233-770, 235-845, 237-562, 238-912, 248-507, 255-721, 257-832, 258-524, 258-788, 263-928, 264-540, 264-826, 265-642, 266-470, 266-527, 270-645, 270-684, 278-556, 281-755, 282-786, 285-856, 289-840, 289-933, 290-787, 292-839, 293-488, 297-589, 297-867, 299-810, 309-824, 312-758, 312-822, 312-841, 312-928, 312-929, 312-934, 313-923, 314-561, 316-440, 316-561, 316-629, 316-658, 316-751, 316-775, 316-788, 316-908, 316-917, 320-933, 325-550, 325-929, 326-693, 326-838, 333-615, 335-853, 337-854, 343-846, 344-540, 348-567, 349-838, 352-872, 352-929, 357-876, 362-731, 366-929, 369-879, 371-821, 386-558, 386-878, 388-638, 394-659, 398-691, 401-623,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	404-912, 406-723, 408-734, 408-737, 411-700, 412-680, 415-726, 418-856, 418-933, 419-737, 422-638, 424-907, 427-911, 431-711, 434-668, 434-677, 434-678, 437-840, 441-671, 441-702, 441-708, 441-731, 444-697, 451-912, 459-739, 461-871, 467-784, 469-726, 476-812, 482-644, 483-753, 485-759, 494-799, 497-847, 499-758, 504-779, 506-741, 506-821, 510-748, 514-690, 516-764, 524-674, 524-811, 526-711, 530-722, 532-781, 541-775, 542-785, 542-801, 542-808, 545-783, 545-789, 547-923, 549-769, 556-818, 557-772, 557-810, 559-854, 564-789, 569-822, 570-839, 575-863, 577-848, 582-731, 594-778, 595-888, 596-837, 606-821, 620-880, 638-881, 638-887, 647-890, 649-863, 654-925, 662-879, 666-803, 666-809, 668-848, 674-929, 701-824, 709-934, 740-835, 928-1165, 928-1191, 931-1174, 932-1196, 944-1142, 949-1025, 949-1158, 961-1166, 967-1159, 985-1188, 988-1145, 1016-1156, 1023-1162, 1028-1175, 1068-1196
86/7509113CB1/ 1419	1-241, 1-1400, 8-287, 11-305, 11-423, 11-425, 14-273, 14-284, 24-265, 24-287, 24-323, 28-305, 31-313, 31-317, 31-352, 31-401, 31-425, 33-192, 33-276, 33-420, 33-425, 34-334, 37-301, 37-665, 38-283, 38-330, 49-254, 49-420, 49-425, 55-358, 57-380, 58-324, 58-335, 58-344, 58-354, 58-373, 68-286, 69-379, 82-381, 83-384, 99-425, 103-425, 147-366, 177-366, 448-1392, 467-698, 504-1397, 505-1393, 547-1397, 569-1284, 600-1393, 603-1107, 603-1140, 604-1089, 604-1393, 609-1088, 610-1187, 610-1326, 615-1028, 624-1088, 632-1393, 634-1089, 635-1129, 636-1089, 637-1089, 638-1106, 638-1146, 639-1236, 649-1060, 652-1083, 653-1340, 654-1393, 666-1083, 672-1089, 673-1106, 691-1181, 691-1405, 697-1405, 698-1393, 703-1328, 716-1295, 720-1236, 744-1235, 751-1404, 755-1317, 755-1351, 760-1264, 761-1265, 763-1060, 764-1089, 764-1393, 765-1089, 765-1107, 765-1332, 775-1236, 789-1393, 789-1416, 794-1414, 822-1343, 830-1350, 831-1365, 831-1393, 843-1096, 843-1343, 849-1395, 855-1318, 857-1363, 871-1326, 871-1393, 878-1360, 889-1276, 904-1397, 905-1035, 905-1236, 905-1343, 905-1350, 909-1355, 910-1362, 937-1354, 969-1332, 969-1350, 974-1361, 984-1286, 995-1400, 1020-1375, 1021-1128,
	1052-1332, 1071-1419, 1122-1400, 1162-1400, 1226-1400, 1249-1400, 1296-1404
87/7509140CB1/ 1200	1-231, 1-295, 1-399, 1-454, 2-1182, 3-430, 6-532, 8-597, 14-277, 14-313, 18-295, 21-342, 21-391, 21-408, 21-604, 23-182, 23-266, 23-364, 23-443, 23-444, 23-468, 23-532, 23-537, 23-547, 23-624, 23-701, 23-750, 23-814, 24-324, 25-675, 26-664, 27-686, 27-712, 27-731, 27-744, 27-745, 27-786, 27-787, 27-795, 27-798, 27-815, 27-819, 28-320, 28-688, 39-433, 39-434, 39-482, 39-516, 40-532, 41-661, 45-540, 47-370, 48-314, 48-325, 48-334, 48-344, 58-276, 59-369, 59-604, 80-661, 84-562, 85-564, 89-433, 89-435, 89-660, 90-604, 93-433, 93-628, 94-648, 94-657, 107-565, 129-660, 137-531, 137-580, 137-593, 137-684, 138-686, 149-661, 166-757, 167-356, 167-517, 167-532, 167-607, 167-634, 167-648, 167-661, 167-717, 177-642, 179-661, 181-769, 194-514, 245-706, 245-716, 245-757, 245-783, 246-717, 249-522, 251-517, 254-1182, 270-757, 277-818, 286-806, 306-394, 310-598, 317-517, 320-526, 324-782, 346-538, 351-630, 357-722, 408-625, 430-743, 580-770, 626-717, 814-1138, 815-917, 815-1164, 815-1189, 815-1193, 815-1200,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	818-1180, 841-1121, 841-1132, 841-1200, 860-1200, 879-1200, 895-1200, 905-1195, 911-1189, 920-1184, 922-1200, 951-1189, 962-1200, 965-1200, 1000-1192, 1015-1189, 1038-1189, 1039-1200, 1108-1191
88/7509223CB1/ 1982	1-945, 604-763, 604-812, 604-836, 604-844, 604-847, 604-855, 604-858, 604-876, 604-884, 604-888, 604-894, 604-923, 604-972, 604-1011, 604-1013, 604-1015, 604-1024, 604-1025, 604-1028, 604-1029, 604-1049, 604-1072, 604-1097, 604-1113, 604-1118, 604-1128, 604-1132, 604-1175, 604-1178, 604-1205, 604-1255, 604-1282, 604-1331, 604-1402, 605-905, 606-1256, 607-1245, 608-872, 608-1483, 609-854, 609-901, 620-825, 620-1009, 620-1014, 620-1015, 620-1097, 620-1124, 621-1113, 622-1242, 626-929, 626-1121, 628-951, 629-895, 629-906, 629-915, 629-925, 629-944, 639-857, 640-950, 640-1185, 653-952, 654-955, 661-1242, 665-1143, 666-1145, 670-1014, 670-1203, 670-1241, 671-1185, 674-1014, 674-1209, 675-1229, 675-1238, 688-1146, 710-1241, 718-937, 718-1112, 718-1161, 718-1174, 718-1265, 719-1267, 730-1242, 747-1338, 748-937, 748-1098, 748-1113, 748-1188, 748-1215, 748-1229, 748-1233, 748-1242, 748-1298, 748-1384, 758-1223, 760-1242, 762-1350, 775-1095, 780-1399, 797-1275, 826-1130, 826-1229, 826-1287, 826-1297, 826-1338, 826-1401, 827-1298, 830-1103, 832-1098, 833-1179, 851-1338, 858-1399, 867-1387, 877-1401, 887-1170,
	891-1179, 898-1098, 901-1107, 901-1346, 905-1363, 909-1107, 927-1119, 932-1211, 938-1396, 938-1401, 989-1206, 1011-1324, 1016-1401, 1033-1122, 1050-1293, 1107-1374, 1118-1297, 1131-1298, 1170-1958, 1207-1298, 1396-1654, 1396-1908, 1396-1915, 1396-1930, 1396-1942, 1399-1872, 1408-1661, 1408-1908, 1414-1960, 1420-1883, 1422-1928, 1428-1982, 1436-1891, 1436-1958, 1443-1925, 1445-1902, 1448-1982, 1454-1841, 1470-1600, 1470-1801, 1470-1908, 1470-1915, 1474-1920, 1475-1927, 1502-1919, 1512-1979, 1517-1977, 1533-1967, 1533-1979, 1534-1897, 1534-1915, 1539-1926, 1549-1851, 1560-1965, 1566-1976, 1569-1965, 1577-1672, 1577-1979, 1585-1940, 1586-1693, 1586-1969, 1587-1914, 1594-1956, 1617-1897, 1617-1908, 1617-1979, 1636-1982, 1655-1979, 1671-1982, 1681-1971, 1687-1965, 1696-1960, 1698-1979, 1727-1965, 1738-1979, 1741-1979, 1776-1968, 1791-1965, 1814-1965, 1815-1979, 1861-1969, 1884-1967
89/7509272CB1/ 1282	1-337, 1-1282, 20-439, 23-842, 23-880, 358-985, 358-1117, 358-1282, 361-958, 362-983, 362-985, 364-985, 388-1282, 389-1282, 424-594, 427-1282, 438-985, 462-720, 501-985, 547-985, 638-985, 660-985, 669-985, 773-984, 795-1072, 830-947, 831-1180, 846-1044
90/7509327CB1/ 1337	1-315, 1-1337, 2-315, 20-364, 23-315, 23-839, 371-603, 478-1316, 579-1316, 625-1304, 626-1316, 640-851, 652-1304, 662-939, 672-1316, 697-814, 698-1307, 698-1316, 699-1316, 713-911, 871-1316, 943-1337, 1138-1316
91/7504677CB1/ 843	1-92, 1-152, 2-843, 3-248, 5-268, 12-299, 14-285, 22-260, 30-280, 38-159, 38-172, 39-316, 40-153, 42-292, 43-250, 46-282, 51-316, 59-318, 60-318, 61-154, 63-173, 66-292, 100-318, 134-318, 181-766, 191-451, 258-455, 321-574, 352-767, 390-784, 395-726, 401-618, 403-784, 409-782, 438-767, 442-784, 444-782, 472-783, 478-784, 606-843, 630-758

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
92/7504534CB1/ 3360	1-249, 1-384, 1-446, 1-538, 1-559, 4-293, 7-293, 7-308, 14-326, 14-585, 14-648, 14-3284, 15-698, 21-572, 23-328, 24-378, 24-637, 24-646, 24-848, 25-265, 25-674, 26-397, 26-589, 33-268, 50-514, 104-348, 105-348, 174-704, 186-415, 186-421, 186-732, 271-583, 324-711, 353-522, 376-646, 422-965, 529-767, 590-841, 636-1253, 650-711, 663-1159, 671-1035, 693-963, 716-1325, 789-1011, 807-1063, 813-1338, 829-1120, 837-1432, 863-1668, 886-1531, 913-1507, 933-1210, 997-1230, 1022-1599, 1044-1335, 1059-1338, 1081-1782, 1111-1369, 1111-1391, 1122-1526, 1124-1579, 1144-1793, 1158-1390, 1195-1456, 1195-1477, 1195-1483, 1206-1831, 1218-1346, 1234-1474, 1239-1800, 1249-1522, 1251-1793, 1257-1534, 1258-1670, 1299-1558, 1305-1426, 1305-2167, 1314-1966, 1315-1632, 1335-1596, 1335-1805, 1336-1589, 1358-1630, 1385-1644, 1391-2023, 1418-1903, 1466-1676, 1466-1688, 1489-1800, 1495-1831, 1507-1776, 1518-1744, 1556-2145, 1653-1891, 1674-1974, 1675-1959, 1682-1917, 1682-1925, 1682-2433, 1700-1945, 1704-2308, 1751-2181, 1769-2304, 1773-2038, 1784-2068, 1853-2120, 1854-2407, 1856-2280, 1857-2400, 1859-2059, 1866-2437, 1889-2167,
	1894-2277, 1908-2171, 1913-2179, 1930-2063, 1947-2204, 1947-2401, 1966-2261, 1969-2591, 1994-2276, 2004-2179, 2005-2351, 2010-2311, 2012-2634, 2019-2270, 2019-2272, 2036-2280, 2062-2395, 2066-2340, 2069-2328, 2084-2374, 2092-2335, 2116-2325, 2156-2504, 2173-2383, 2173-2670, 2205-2440, 2216-2666, 2226-2372, 2227-2423, 2245-2637, 2251-2660, 2282-2520, 2334-2574, 2364-2473, 2381-2724, 2384-2666, 2388-2948, 2391-2644, 2413-2952, 2432-2666, 2444-2666, 2460-2666, 2471-2751, 2471-2913, 2474-2778, 2474-2956, 2487-2972, 2490-2666, 2510-2960, 2587-2860, 2587-2972, 2629-2701, 2632-2952, 2665-2950, 2665-2960, 2665-2961, 2665-3072, 2665-3146, 2667-2923, 2671-2960, 2674-2945, 2685-2951, 2697-2972, 2718-2951, 2719-2997, 2735-2974, 2735-2987, 2753-2991, 2762-3083, 2763-3284, 2773-2950, 2791-3287, 2812-3339, 2840-3134, 2849-2994, 2864-2960, 2887-2952, 2927-3360, 2939-3181, 2953-3213, 2954-3360, 2959-3284, 2960-3167, 2961-3251, 2975-3214, 2975-3231, 2976-3272, 2980-3186, 2995-3258, 2995-3297, 3002-3236, 3003-3288, 3027-3278, 3041-3284, 3161-3282, 3163-3284

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
93/7507771CB1/ 13855	1-578, 145-577, 158-713, 158-4257, 196-849, 205-590, 206-585, 228-929, 262-601, 300-435, 435-1092, 503-888, 505-668, 576-946, 692-1048, 745-1155, 890-1336, 1194-1815, 1233-1944, 1252-1727, 1501-1875, 1563-1665, 1672-2281, 1672-2302, 1672-2335, 1680-2379, 1688-2027, 1833-2476, 1843-2348, 1884-2089, 1921-2476, 1944-2425, 1947-2055, 1947-2152, 1947-2477, 1960-2153, 2249-2677, 2435-2666, 2483-2628, 2483-2948, 2483-2996, 2483-3032, 2483-3047, 2483-3052, 2493-3052, 2499-2628, 2501-3139, 2502-3050, 2505-2750, 2571-2960, 2617-3050, 2622-3031, 2640-3055, 2661-2839, 2673-3356, 2694-3198, 2721-2778, 2813-3069, 2813-3397, 2813-3462, 2903-3050, 2907-3424, 2955-3108, 2995-3350, 2995-3352, 3023-3424, 3036-3508, 3042-3166, 3098-3812, 3188-3824, 3312-3705, 3655-3831, 3661-4101, 3661-13855, 3754-4027, 3761-3959, 3761-4098, 3839-4341, 3848-4444, 3984-4485, 4005-4490, 4016-4447, 4107-4354, 4107-4632, 4192-4472, 4195-4428, 4241-4482, 4257-4664, 4262-4845, 4344-4931, 4344-4983, 4455-5009, 4470-5045, 4551-4846, 4551-5196, 4566-4937, 4571-4937, 4635-4896, 4635-5204, 4680-5046, 4732-5337, 4756-4859, 4757-5441,
	4757-13855, 4759-5040, 4798-5375, 4835-5304, 4848-5426, 4859-5246, 4941-5237, 4941-5457, 4956-5556, 4991-5455, 5053-5387, 5091-5811, 5158-6011, 5177-5728, 5194-5899, 5202-13855, 5287-5320, 5364-5899, 5437-5899, 5566-5916, 5574-6147, 5678-6372, 5681-5923, 5684-6147, 6218-6495, 6218-6542, 6218-6771, 6218-13855, 6229-6758, 6231-6819, 6267-6830, 6376-6790, 6376-6812, 6379-6813, 6391-6818, 6409-6818, 6411-6677, 6419-7067, 6419-7078, 6429-6818, 6429-6819, 6433-6819, 6434-6942, 6438-6784, 6438-6819, 6438-7206, 6441-6819, 6480-7082, 6489-6818, 6495-6920, 6505-6937, 6507-6848, 6518-6819, 6600-6831, 6601-6818, 6606-7070, 6606-7081, 6631-7258, 6635-6703, 6637-7422, 6643-7316, 6666-6978, 6724-7365, 6746-7290, 6801-7399, 6802-7324, 6804-7324, 6893-7582, 6896-7225, 6902-7485, 6921-7354, 6931-7393, 6943-7470, 6952-7483, 6963-7538, 7021-7292, 7032-7564, 7034-7371, 7034-7503, 7035-7226, 7049-7406, 7072-7686, 7075-7358, 7090-7564, 7112-7564, 7117-7253, 7130-7284, 7164-7339, 7164-7737, 7179-7674, 7234-7842, 7322-7753, 7340-13855, 7348-7753, 7414-7753, 7460-7911, 7465-7937, 7465-7981, 7484-7753,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	7498-7984, 7506-7748, 7511-7841, 7513-8366, 7537-7753, 7538-8080, 7550-7988, 7553-7964, 7565-8375, 7589-7753, 7598-7911, 7598-7986, 7607-7753, 7656-7753, 7667-7750, 7682-8153, 7693-8312, 7693-8393, 7701-8296, 7763-8604, 7847-8520, 7936-8538, 7936-8554, 7955-8479, 7962-8565, 8001-13855, 8004-8379, 8064-8418, 8071-8397, 8093-8664, 8103-8377, 8107-8368, 8111-8738, 8132-8594, 8134-8594, 8140-8398, 8144-8582, 8169-8586, 8243-8500, 8264-8803, 8317-8835, 8317-8986, 8341-8936, 8347-8894, 8400-8962, 8448-8997, 8643-8955, 8653-8999, 8724-9447, 8745-8993, 8779-9320, 8779-9339, 8781-9059, 8781-9243, 8781-9255, 8782-9269, 8860-9444, 8962-13855, 8964-9788, 9008-9266, 9131-9939, 9199-9623, 9215-9939, 9238-9812, 9244-9871, 9266-9838, 9341-9623, 9361-9752, 9399-10073, 9404-9812, 9414-9805, 9428-9836, 9429-9972, 9453-9836, 9456-9899, 9465-9830, 9473-9809, 9483-9796, 9508-9868, 9525-9829, 9549-9878, 9560-9618, 9560-10097, 9560-13855, 9561-10356, 9561-10383, 9574-9922, 9601-10262, 9693-10388, 9725-10428, 9811-10327, 9813-10620, 9870-10310, 9907-10425, 9939-10396, 9964-10579, 9990-10386, 9991-10477, 10027-10371,
	10105-10659, 10114-10723, 10114-13855, 10131-10918, 10150-10460, 10166-10509, 10170-10278, 10177-10838, 10180-10792, 10197-10909, 10201-10677, 10212-10703, 10216-10890, 10237-10810, 10244-10815, 10296-10847, 10316-10846, 10317-10877, 10354-10907, 10367-10944, 10371-10654, 10378-10702, 10378-10844, 10387-10779, 10391-10751, 10406-10804, 10407-10917, 10410-10714, 10413-10747, 10424-10734, 10449-10965, 10460-10934, 10495-10755, 10495-10765, 10496-10928, 10510-10968, 10516-10712, 10516-10869, 10524-10921, 10539-10961, 10541-10916, 10541-11194, 10552-10741, 10552-11145, 10568-10714, 10569-10714, 10572-10921, 10572-10934, 10572-10938, 10577-10775, 10577-10840, 10577-10850, 10577-10856, 10577-10917, 10577-10921, 10577-10931, 10577-10934, 10579-10931, 10603-11307, 10603-11314, 10603-11406, 10603-13855, 10606-10873, 10624-10934, 10637-11141, 10650-11089, 10656-10931, 10711-11395, 10754-11412, 10758-11411, 10838-11330, 10873-11518, 10897-11616, 10909-11563, 10910-10972, 10941-11538, 10944-11614, 10951-11410, 10959-11414, 10959-11462, 10959-11497, 10959-11499, 10959-11500, 10959-11501, 10959-11510,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	10962-11448, 10981-11595, 10992-11327, 10992-11486, 11008-11634, 11019-11445, 11023-11694, 11026-11645, 11031-11662, 11054-11515, 11055-11310, 11064-11665, 11077-11352, 11083-11698, 11108-11378, 11121-11785, 11126-11344, 11143-11364, 11153-11364, 11153-11523, 11159-11565, 11163-11824, 11178-11789, 11180-11792, 11225-11885, 11225-12985, 11265-11957, 11333-12169, 11347-11919, 11351-11818, 11363-11994, 11455-12161, 11457-11838, 11467-12102, 11517-12034, 11629-12169, 11669-12191, 11697-12308, 11722-12277, 11737-12226, 11776-12263, 12000-12280, 12097-12753, 12207-12811, 12246-12860, 12248-12804, 12255-12768, 12259-12765, 12273-12841, 12287-12759, 12288-12863, 12292-12804, 12294-12838, 12297-12794, 12312-12930, 12316-12908, 12328-12712, 12330-12995, 12333-12855, 12341-12925, 12343-12836, 12343-12886, 12345-12865, 12370-13190, 12379-12821, 12389-12763, 12392-13204, 12393-12958, 12409-12725, 12418-13049, 12425-12945, 12432-13216, 12432-13220, 12445-12775, 12456-12794, 12457-13026, 12472-13051, 12476-12935, 12495-12797, 12496-13081, 12524-12806, 12525-13036, 12525-13041, 12527-12985,
	12528-13335, 12532-13031, 12559-13059, 12559-13245, 12561-12792, 12561-12816, 12565-12857, 12571-12985, 12577-13413, 12582-13174, 12587-12876, 12590-13142, 12598-12825, 12601-12885, 12603-13224, 12604-12879, 12608-12891, 12608-13233, 12616-13244, 12617-12877, 12622-12889, 12622-13116, 12624-12794, 12625-12725, 12635-13022, 12641-13059, 12643-12870, 12647-12970, 12661-12809, 12663-13446, 12676-13313, 12685-12976, 12696-12915, 12696-13272, 12724-13000, 12728-13410, 12742-13401, 12752-13038, 12758-13276, 12760-13490, 12775-13486, 12777-13456, 12780-13421, 12783-13376, 12794-13268, 12803-13461, 12819-13141, 12819-13216, 12819-13287, 12821-13530, 12832-13359, 12845-13524, 12860-13095, 12876-13543, 12880-13153, 12880-13327, 12880-13383, 12891-13416, 12892-13337, 12905-13460, 12908-13370, 12917-13591, 12924-13795, 12928-13203, 12930-13159, 12935-13174, 12935-13226, 12940-13193, 12940-13216, 12942-13213, 12951-13652, 12957-13191, 12959-13446, 12962-13220, 12967-13530, 12973-13221, 12973-13237, 12973-13397, 12973-13398, 12973-13518, 12974-13216, 12975-13192, 12981-13237, 12983-13625, 12983-13729,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	12986-13235, 12988-13202, 12990-13188, 12990-13481, 12992-13571, 12993-13698, 12995-13311, 12996-13284, 12996-13507, 12997-13614, 12998-13258, 13003-13546, 13009-13264, 13009-13695, 13012-13253, 13012-13314, 13012-13696, 13015-13467, 13015-13617, 13016-13195, 13024-13284, 13035-13559, 13039-13292, 13041-13773, 13051-13311, 13051-13715, 13054-13853, 13056-13540, 13059-13487, 13059-13502, 13059-13512, 13060-13361, 13060-13506, 13064-13211, 13067-13800, 13067-13855, 13068-13218, 13068-13629, 13071-13300, 13072-13330, 13073-13697, 13076-13324, 13077-13360, 13077-13536, 13080-13255, 13080-13355, 13082-13267, 13087-13696, 13095-13368, 13095-13377, 13095-13380, 13097-13346, 13100-13551, 13105-13359, 13105-13363, 13106-13262, 13107-13343, 13108-13613, 13116-13659, 13122-13840, 13130-13377, 13133-13373, 13133-13386, 13137-13655, 13147-13501, 13149-13417, 13159-13427, 13159-13466, 13164-13412, 13165-13249, 13166-13288, 13168-13465, 13171-13448, 13179-13216, 13202-13244, 13206-13340, 13214-13249
94/7504732CB1/ 1554	1-177, 2-1552, 25-177, 43-515, 43-517, 47-177, 56-254, 60-177, 70-177, 74-177, 89-177, 93-177, 102-177, 117-175, 117-177, 139-355, 451-1040, 460-979, 460-1089, 472-1256, 478-655, 478-715, 478-1118, 478-1180, 504-942, 504-971, 526-1301, 570-1113, 583-1406, 648-1248, 687-1389, 720-1554, 748-1366, 751-1553, 768-1553, 780-1554, 786-1390, 794-1324, 794-1553, 812-1551, 818-1471, 828-1386, 881-1140, 885-1547, 916-1150, 929-1551, 930-1426, 943-1538, 959-1181, 959-1315, 959-1494, 959-1507, 959-1509, 959-1510, 959-1512, 959-1515, 959-1530, 959-1535, 959-1541, 959-1546, 959-1554, 987-1227, 991-1323, 991-1324, 1014-1472, 1016-1334, 1028-1471, 1092-1536, 1147-1408, 1163-1398
95/950917CB1/ 2142	1-752, 92-752, 104-752, 109-752, 118-752, 161-750, 161-858, 166-752, 201-902, 211-752, 228-904, 237-752, 240-752, 254-752, 273-907, 283-752, 287-858, 288-752, 310-752, 315-752, 321-752, 322-752, 334-752, 354-752, 359-752, 368-904, 429-907, 448-752, 456-752, 584-907, 635-1818, 700-907, 704-907, 705-907, 706-907, 709-907, 727-907, 783-907, 853-907, 1736-1900, 1736-1903, 1736-1956, 1736-1996, 1736-2023, 1736-2031, 1736-2077, 1736-2142, 1746-1902, 1762-2031
96/7459720CB1/ 2637	1-612, 32-426, 32-437, 32-474, 32-673, 32-740, 32-790, 32-818, 32-840, 33-612, 33-789, 33-806, 61-773, 197-967, 269-869, 276-2543, 285-844, 319-856, 319-967, 320-1130, 364-1030, 386-967, 390-989, 422-991, 439-906, 444-945, 449-1139, 450-1030, 467-992, 493-1024, 497-963, 516-1011, 521-1139, 533-955, 535-1030, 542-1008, 542-1248, 556-1221, 609-1021, 613-1128, 644-1171, 644-1285, 671-1237, 683-1233, 691-1130, 693-1404, 698-1220, 710-1190, 715-1265, 715-1372, 731-1204, 732-1206, 735-1260, 739-1324, 742-1459, 746-1387, 777-1348, 816-1504, 820-1374, 828-1242, 828-1515, 844-1551, 855-1120, 972-1027, 1046-1128, 1046-1199, 1049-1094, 1051-1724, 1080-1690, 1106-1690, 1126-1739, 1156-1930, 1206-1992, 1326-2003, 1329-1921, 1372-1957, 1408-2094, 1424-2100, 1455-2106, 1466-2171, 1596-2205, 1596-2242, 1671-2354, 1739-2323, 1811-2637, 1831-2468, 1839-2525, 1839-2630, 1839-2637, 1863-2519, 1880-2507, 1880-2541, 1905-2486, 1940-2581, 1944-2593

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
97/7503300CB1/ 2447	1-290, 1-2414, 56-856, 58-368, 58-500, 81-476, 86-423, 87-420, 97-226, 97-244, 97-258, 97-349, 97-360, 97-502, 97-557, 97-574, 97-589, 97-593, 97-611, 97-692, 97-720, 100-535, 112-770, 130-256, 133-379, 134-427, 137-698, 167-427, 175-458, 264-485, 297-781, 297-897, 438-1000, 455-1003, 640-1206, 758-1126, 758-1190, 917-1181, 988-1206, 1029-1206, 1109-1467, 1109-1764, 1109-1899, 1206-1509, 1213-1542, 1244-1496, 1253-1709, 1253-1805, 1253-1820, 1253-1882, 1256-1509, 1256-1547, 1280-1571, 1309-1825, 1310-1979, 1348-1648, 1366-1576, 1500-2005, 1504-1771, 1514-2098, 1515-1748, 1534-1820, 1537-1781, 1553-2011, 1568-2129, 1577-2390, 1581-1999, 1582-1763, 1582-2057, 1586-2237, 1604-1835, 1624-2270, 1629-2210, 1631-2332, 1692-2173, 1693-2245, 1693-2299, 1702-2242, 1705-2313, 1711-2134, 1716-1947, 1769-2222, 1777-2142, 1779-2326, 1783-2231, 1784-2392, 1785-2358, 1790-2339, 1805-2025, 1826-2381, 1829-2099, 1829-2232, 1836-2276, 1838-2082, 1839-2376, 1855-2405, 1861-2134, 1895-2148, 1913-2389, 1913-2414, 1925-2170, 1925-2216, 1925-2229, 1925-2233, 1930-2173, 1931-2280, 1935-2399, 1938-2396, 1939-2186, 1939-2382, 1939-2399,
	1944-2134, 1948-2210, 1949-2238, 1951-2402, 1961-2401, 1962-2395, 1971-2399, 1977-2336, 1988-2285, 1988-2399, 1993-2305, 1995-2399, 1998-2403, 2002-2313, 2004-2393, 2005-2432, 2008-2162, 2009-2395, 2010-2172, 2012-2401, 2025-2399, 2029-2397, 2030-2164, 2036-2274, 2036-2399, 2071-2395, 2085-2395, 2096-2305, 2116-2384, 2122-2395, 2132-2402, 2165-2380, 2188-2365, 2200-2447, 2201-2400, 2202-2398, 2207-2379, 2207-2413, 2207-2425, 2216-2425, 2240-2356, 2240-2384, 2240-2395, 2253-2425, 2264-2397, 2284-2355
98/7503334CB1/ 2384	1-2035, 2-787, 101-215, 101-216, 101-382, 101-404, 101-412, 101-588, 101-601, 101-624, 101-660, 101-668, 111-752, 116-787, 118-404, 122-760, 134-403, 134-404, 137-705, 146-729, 149-440, 150-466, 154-738, 156-726, 160-748, 163-804, 166-853, 171-771, 174-943, 179-829, 189-404, 196-399, 196-952, 197-887, 200-404, 203-807, 203-832, 206-404, 212-760, 213-697, 213-698, 213-699, 216-951, 218-297, 218-708, 221-698, 229-476, 229-616, 234-950, 236-508, 238-717, 242-952, 249-754, 253-487, 254-404, 254-852, 256-404, 262-867, 270-950, 276-602, 278-837, 295-895, 296-949, 296-953, 297-917, 298-947, 305-949, 313-916, 319-825, 329-404, 330-589, 333-404, 346-600, 360-952, 365-902, 369-1030, 373-677, 382-999, 384-1059, 392-897, 399-656, 404-960, 406-659, 408-1057, 412-851, 415-978, 416-799, 419-1017, 439-992, 448-672, 449-1042, 453-1078, 454-1035, 456-1073, 457-1073, 461-779, 472-999, 479-960, 488-759, 503-947, 503-951, 503-952, 523-1103, 530-762, 533-1469, 538-816, 539-824, 544-762, 545-1149, 549-1193, 550-807, 551-821, 563-833, 569-1202,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	571-942, 572-812, 572-858, 576-822, 596-895, 601-1173, 605-803, 605-1048, 605-1112, 623-951, 623-952, 625-1263, 628-1032, 645-916, 645-1159, 647-919, 647-920, 653-1149, 687-1343, 687-1344, 716-951, 717-969, 753-1032, 757-953, 758-1020, 769-925, 775-1071, 775-1317, 777-1394, 777-1452, 800-1133, 821-1446, 822-1054, 828-1380, 833-1108, 847-1438, 854-1752, 859-1521, 863-1441, 865-1753, 867-951, 869-944, 869-951, 869-952, 869-953, 869-1110, 872-1128, 878-1753, 888-951, 890-1753, 903-1753, 905-1153, 910-1165, 920-1032, 925-1571, 926-1589, 932-1565, 939-1293, 940-1564, 941-1589, 948-1263, 953-1188, 957-1087, 967-1608, 973-1753, 978-1532, 981-1753, 983-1753, 984-1753, 993-1597, 996-1753, 998-1460, 1008-1636, 1013-1461, 1014-1160, 1018-1260, 1023-1599, 1024-1212, 1025-1753, 1043-1754, 1045-1713, 1056-1304, 1056-1309, 1072-1623, 1074-1557, 1075-1702, 1084-1750, 1087-1717, 1087-1757, 1092-1305, 1099-1641, 1100-1365, 1102-1363, 1107-2011, 1110-1554, 1115-1365, 1115-1673, 1120-1398, 1120-1590, 1121-1790, 1126-1507, 1129-1613, 1147-1608, 1155-1788, 1158-1570, 1160-1600, 1166-1967, 1170-1961, 1172-1436, 1172-1788,
	1191-1350, 1197-1841, 1197-1977, 1214-1608, 1220-1492, 1220-1509, 1220-1739, 1226-1641, 1256-1496, 1256-1784, 1257-1899, 1279-1558, 1290-2001, 1304-1455, 1304-1871, 1305-1570, 1320-1626, 1321-1994, 1326-1533, 1328-1595, 1330-2004, 1340-1631, 1340-1976, 1341-1596, 1346-2015, 1351-1981, 1365-1665, 1378-1572, 1381-2025, 1399-1986, 1403-1973, 1408-2077, 1421-2025, 1430-2016, 1433-2057, 1443-2022, 1449-2033, 1460-1748, 1461-2053, 1466-1986, 1466-2049, 1467-1700, 1467-2029, 1476-1710, 1481-1744, 1483-1704, 1483-1766, 1489-1752, 1489-2034, 1491-2024, 1492-1749, 1492-2089, 1493-2050, 1500-1988, 1502-2064, 1509-1737, 1513-1770, 1514-1761, 1517-1735, 1519-2031, 1519-2032, 1521-1711, 1524-1829, 1524-2015, 1525-2028, 1527-2053, 1528-2035, 1535-1678, 1535-1755, 1537-2086, 1544-1799, 1555-2035, 1563-2051, 1565-2035, 1567-2044, 1568-2041, 1582-2040, 1583-2384, 1592-2082, 1594-2047, 1595-1840, 1595-2041, 1596-1830, 1597-1867, 1599-2031, 1601-1843, 1603-2091, 1604-1922, 1611-2002, 1616-2040, 1623-1871, 1625-2033, 1626-1905, 1626-1922, 1626-2035, 1626-2053, 1627-2028, 1630-2040, 1632-1873, 1639-2035, 1642-1899, 1648-1890,
	1659-1861, 1659-2038, 1661-2034, 1664-2091, 1670-1796, 1677-1942, 1678-2045, 1684-2038, 1692-2054, 1699-2093, 1703-2033, 1709-1960, 1709-2040, 1713-2035, 1719-2002, 1721-2054, 1722-2054, 1723-1986, 1726-2040, 1728-2035, 1746-2035, 1751-1953, 1751-2027, 1751-2052, 1764-1813, 1765-2020, 1770-2046, 1798-2034, 1814-2055, 1898-2034
99/7503341CB1/ 778	1-489, 17-577, 20-243, 20-259, 20-586, 21-234, 21-274, 21-465, 21-578, 21-579, 22-579, 24-465, 25-531, 26-123, 26-584, 26-778, 95-194, 276-400, 284-586, 330-586, 336-587, 337-586, 346-586, 359-586, 360-586, 475-769, 575-760, 575-767, 575-768, 575-778, 576-754, 583-768, 606-778, 613-778, 637-765, 640-765, 642-778, 684-768, 692-768, 693-774, 699-771

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
100/7509936CB1/ 3705	1-110, 7-3705, 40-319, 162-451, 552-1304, 580-1194, 614-945, 614-1131, 614-1170, 614-1190, 614-1193, 656-1227, 769-1063, 870-1029, 875-1029, 1035-1495, 1052-1245, 1055-1078, 1190-1365, 1211-1944, 1285-1913, 1458-1746, 1608-1890, 1608-2166, 1639-2425, 1642-2023, 1642-2025, 1790-2602, 1943-2519, 1943-2535, 2173-2834, 2183-2812, 2197-2479, 2197-2483, 2200-2705, 2222-2869, 2256-2850, 2296-2454, 2299-2444, 2316-2608, 2368-3157, 2389-3217, 2558-3217, 2589-3235, 2847-3656, 2945-3322, 2947-3407, 2975-3178, 3009-3434, 3020-3456, 3050-3612, 3066-3246, 3095-3456, 3144-3500, 3154-3576, 3157-3405, 3157-3408, 3173-3450, 3175-3412, 3186-3462, 3204-3622, 3227-3497, 3263-3501, 3299-3702, 3305-3527, 3310-3414, 3317-3673, 3326-3699, 3371-3617, 3386-3620, 3450-3690, 3481-3617, 3486-3617
101/7509986CB1/ 3872	1-219, 1-561, 44-294, 46-553, 69-736, 109-613, 131-757, 190-485, 190-656, 230-478, 233-736, 242-400, 253-516, 253-677, 358-1003, 367-410, 367-412, 367-419, 367-437, 367-438, 367-443, 367-445, 367-452, 367-470, 367-471, 367-476, 381-410, 385-410, 388-443, 393-443, 393-470, 395-443, 395-635, 400-509, 424-692, 426-476, 428-476, 428-509, 433-509, 451-476, 454-1063, 459-509, 459-585, 461-509, 461-704, 466-509, 484-638, 492-1016, 492-1038, 577-1162, 590-776, 640-1146, 645-900, 645-1078, 645-1089, 652-915, 655-915, 659-1472, 665-1311, 666-933, 666-1083, 681-826, 682-1209, 701-1028, 738-1260, 740-1143, 740-1348, 741-1050, 741-1197, 747-1234, 749-1132, 754-1038, 754-1333, 764-1348, 766-1004, 779-1023, 785-1307, 786-1325, 811-1059, 811-1078, 818-1011, 827-1095, 859-1322, 863-1343, 881-1394, 889-1305, 904-1165, 906-1106, 911-1578, 916-1363, 927-1177, 927-1195, 927-1345, 927-1460, 927-1506, 936-1183, 939-1168, 941-1390, 943-1364, 945-1185, 946-1367, 950-1190, 950-1328, 958-1170, 960-1425, 961-1579, 977-1603, 979-1364, 985-1227, 985-1247, 987-1507, 990-1101, 1010-1476, 1014-1235, 1026-1394, 1029-1646, 1035-1287,
	1037-1514, 1051-1519, 1052-1688, 1057-1289, 1057-1461, 1067-1715, 1088-1331, 1091-1295, 1112-1699, 1113-1699, 1121-1371, 1122-1615, 1142-1432, 1146-1730, 1159-1493, 1162-1401, 1166-1270, 1168-1397, 1169-1345, 1189-1308, 1205-1405, 1205-1416, 1221-1400, 1232-1511, 1233-1507, 1275-1531, 1284-1346, 1287-1492, 1311-1605, 1328-1853, 1365-1897, 1373-1859, 1443-1817, 1490-1687, 1493-1737, 1495-1795, 1526-1726, 1828-2043, 1833-2030, 1879-2006, 1895-2132, 1895-2567, 1895-2644, 1902-2055, 1905-2421, 1922-2211, 1937-2503, 1964-2235, 1965-2227, 2025-2281, 2068-2377, 2072-2295, 2079-2365, 2505-2788, 2506-3063, 2756-3010, 2769-3147, 2779-2990, 2779-3097, 2801-3011, 2805-3371, 2808-3309, 2808-3473, 2818-3042, 2829-3676, 2841-3080, 2843-3095, 2849-3088, 2849-3103, 2849-3107, 2851-3044, 2854-3162, 2858-3122, 2867-3121, 2872-3108, 2874-3109, 2874-3119, 2877-3141, 2877-3143, 2878-3105, 2878-3151, 2878-3254, 2882-3147, 2893-3025, 2895-3163, 2895-3651, 2903-3142, 2904-3198, 2905-3492, 2906-3165, 2906-3188, 2909-3175, 2911-3140, 2911-3170, 2911-3174, 2911-3181, 2912-3117, 2912-3137, 2912-3151, 2912-3170, 2914-3173,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	2915-3089, 2916-3192, 2916-3206, 2920-3675, 2921-3182, 2923-3205, 2924-3186, 2924-3272, 2924-3430, 2924-3686, 2926-3430, 2931-3013, 2934-3188, 2934-3683, 2948-3281, 2950-3170, 2950-3215, 2950-3228, 2950-3235, 2950-3681, 2951-3243, 2957-3184, 2957-3236, 2959-3161, 2961-3258, 2969-3201, 2969-3263, 2972-3225, 2972-3242, 2974-3180, 2974-3213, 2975-3276, 2976-3215, 2976-3220, 2976-3272, 2976-3274, 2977-3197, 2977-3226, 2979-3464, 2983-3237, 2987-3730, 2989-3275, 2990-3181, 2990-3239, 2990-3242, 2993-3514, 2995-3274, 2995-3287, 2995-3299, 2995-3526, 2996-3217, 2996-3250, 2996-3254, 2996-3260, 2997-3727, 2998-3225, 2998-3246, 2999-3290, 3003-3101, 3004-3664, 3005-3252, 3005-3279, 3006-3221, 3006-3259, 3007-3268, 3008-3243, 3009-3336, 3015-3458, 3021-3256, 3022-3259, 3028-3140, 3029-3134, 3032-3309, 3039-3287, 3041-3730, 3046-3578, 3049-3167, 3050-3310, 3051-3331, 3061-3504, 3063-3744, 3064-3588, 3064-3727, 3067-3730, 3068-3326, 3069-3337, 3069-3358, 3070-3441, 3075-3333, 3079-3331, 3082-3535, 3088-3329, 3089-3321, 3089-3387, 3094-3328, 3096-3310, 3096-3342, 3100-3354, 3103-3331, 3105-3493, 3110-3564, 3116-3686,
	3122-3367, 3124-3453, 3129-3719, 3131-3346, 3132-3332, 3132-3367, 3132-3404, 3134-3397, 3137-3380, 3143-3382, 3150-3386, 3155-3431, 3155-3473, 3155-3593, 3163-3688, 3166-3408, 3166-3442, 3166-3479, 3170-3689, 3173-3390, 3176-3690, 3178-3763, 3181-3687, 3190-3678, 3191-3457, 3192-3777, 3193-3394, 3193-3583, 3195-3688, 3201-3315, 3209-3451, 3217-3310, 3217-3731, 3222-3500, 3223-3388, 3223-3670, 3228-3747, 3232-3761, 3237-3343, 3241-3281, 3241-3282, 3247-3739, 3253-3688, 3259-3382, 3270-3535, 3280-3744, 3281-3578, 3283-3486, 3283-3536, 3283-3833, 3293-3551, 3293-3561, 3294-3687, 3298-3611, 3299-3605, 3304-3746, 3305-3747, 3305-3749, 3305-3755, 3307-3551, 3310-3573, 3311-3531, 3312-3490, 3312-3572, 3315-3756, 3316-3766, 3318-3511, 3318-3724, 3319-3580, 3319-3583, 3320-3581, 3324-3560, 3335-3583, 3336-3517, 3336-3526, 3336-3601, 3341-3585, 3342-3631, 3349-3589, 3352-3721, 3356-3761, 3357-3735, 3358-3726, 3360-3627, 3360-3628, 3360-3647, 3360-3725, 3360-3726, 3360-3727, 3360-3729, 3360-3731, 3360-3734, 3361-3624, 3361-3628, 3361-3645, 3361-3653, 3361-3655, 3361-3726, 3363-3610, 3363-3619, 3363-3754, 3364-3676,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	3364-3740, 3365-3610, 3367-3626, 3368-3726, 3369-3546, 3372-3496, 3372-3528, 3372-3726, 3373-3727, 3373-3733, 3374-3673, 3374-3735, 3375-3621, 3375-3655, 3376-3732, 3380-3727, 3381-3726, 3394-3650, 3396-3735, 3399-3731, 3402-3547, 3402-3659, 3402-3667, 3402-3671, 3407-3611, 3407-3691, 3407-3727, 3408-3612, 3408-3613, 3408-3650, 3408-3658, 3408-3683, 3408-3727, 3410-3732, 3412-3730, 3413-3726, 3417-3669, 3417-3739, 3418-3667, 3419-3668, 3419-3707, 3419-3716, 3420-3673, 3422-3694, 3423-3751, 3425-3631, 3425-3651, 3425-3655, 3429-3687, 3429-3696, 3435-3726, 3436-3710, 3437-3676, 3437-3727, 3438-3726, 3438-3757, 3441-3727, 3442-3726, 3442-3735, 3445-3726, 3446-3587, 3446-3700, 3446-3727, 3447-3726, 3447-3741, 3449-3650, 3449-3726, 3450-3621, 3454-3712, 3458-3729, 3459-3705, 3460-3717, 3469-3739, 3475-3743, 3476-3740, 3479-3726, 3479-3762, 3483-3728, 3487-3745, 3491-3590, 3494-3738, 3497-3742, 3508-3767, 3509-3726, 3520-3741, 3521-3733, 3521-3774, 3524-3872, 3527-3749, 3530-3741, 3539-3743, 3539-3744, 3544-3771, 3544-3798, 3546-3745, 3547-3726, 3549-3726, 3550-3740, 3550-3741, 3552-3726, 3552-3768, 3555-3763, 3557-3735, 3575-3734, 3578-3726, 3593-3801, 3598-3726, 3600-3794, 3603-3787, 3604-3750, 3604-3769
102/7510010CB1/ 2545	1-525, 21-662, 23-656, 32-285, 33-277, 33-701, 34-285, 35-319, 35-323, 36-330, 36-892, 37-335, 40-170, 40-277, 40-550, 48-283, 50-323, 136-212, 168-323, 170-408, 182-434, 197-680, 241-753, 306-907, 318-464, 319-595, 329-903, 329-990, 329-1002, 329-1025, 329-1051, 329-1114, 329-1127, 329-1202, 332-903, 366-627, 369-1143, 390-974, 431-651, 513-593, 513-642, 513-680, 513-794, 513-889, 513-892, 513-908, 513-1004, 519-969, 606-716, 670-829, 771-900, 1102-1729, 1127-1374, 1127-1399, 1127-1603, 1146-1759, 1171-1596, 1180-1589, 1182-1390, 1200-1479, 1214-1988, 1225-1777, 1232-1642, 1233-1481, 1239-1675, 1246-1988, 1248-1692, 1252-1657, 1287-1898, 1296-1988, 1309-1749, 1320-1905, 1321-1792, 1321-1918, 1350-1628, 1351-1919, 1373-1645, 1377-1785, 1378-1947, 1379-1691, 1384-1949, 1389-1988, 1396-1821, 1400-1611, 1402-1989, 1406-1658, 1419-1850, 1426-1517, 1427-1856, 1427-1859, 1439-1661, 1444-1988, 1446-1988, 1451-1865, 1467-1698, 1470-1988, 1474-1695, 1477-1616, 1477-1988, 1479-1666, 1479-1733, 1492-1655, 1494-1789, 1496-1912, 1511-1925, 1520-1988, 1521-1777, 1524-1767, 1531-1788, 1540-1827, 1543-1971, 1544-1821,
	1544-1988, 1558-1803, 1558-1870, 1560-1988, 1565-1874, 1565-1988, 1579-1842, 1593-1838, 1598-1988, 1605-1988, 1610-1988, 1615-1988, 1616-1901, 1622-1988, 1627-1849, 1627-1988, 1637-1937, 1642-2299, 1645-1988, 1649-1988, 1651-1902, 1664-1916, 1678-1919, 1678-1924, 1683-1964, 1692-1971, 1705-1963, 1709-1994, 1710-1988, 1717-1988, 1732-1988, 1734-1918, 1737-1988, 1746-1988, 1764-1988, 1775-2410, 1808-1988, 1812-1988, 1817-1988, 1818-1988, 1832-1988, 1857-2150, 1877-1988, 1888-1988, 1899-1988, 1903-1988, 1956-1988, 2001-2493, 2083-2545, 2087-2337, 2090-2538, 2091-2545, 2092-2355, 2094-2287, 2097-2545, 2099-2506, 2110-2538, 2117-2545, 2120-2323, 2138-2413, 2139-2359, 2148-2538, 2154-2408, 2160-2277, 2163-2435, 2167-2538, 2175-2499, 2175-2545, 2180-2545, 2185-2540, 2212-2339, 2213-2538, 2216-2496, 2221-2505, 2250-2538, 2253-2520, 2292-2518, 2322-2538, 2328-2538, 2340-2492, 2340-2493, 2340-2500, 2340-2538, 2340-2543, 2408-2540, 2464-2538

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
103/7510056CB1/ 3273	1-242, 1-3273, 257-673, 295-828, 295-839, 295-840, 314-735, 318-696, 394-762, 429-909, 551-1040, 668-1352, 686-921, 741-1360, 753-1061, 771-1341, 787-1124, 789-1428, 833-1249, 849-955, 868-1669, 870-1420, 973-1500, 974-1441, 976-1270, 976-1496, 977-1485, 979-1632, 984-1596, 989-1218, 994-1281, 997-1197, 998-1553, 1002-1291, 1023-1286, 1028-1293, 1037-1692, 1046-1503, 1047-1619, 1057-1499, 1061-1338, 1066-1313, 1068-1295, 1068-1666, 1070-1530, 1074-1530, 1075-1402, 1080-1718, 1091-1652, 1100-1372, 1112-1368, 1112-1572, 1116-1368, 1121-1628, 1128-1655, 1128-1818, 1136-1644, 1142-1608, 1148-1788, 1151-1256, 1151-1633, 1151-1665, 1154-1315, 1154-1375, 1159-1713, 1167-1604, 1182-1736, 1183-1403, 1183-1667, 1189-1849, 1192-1461, 1194-1733, 1199-1487, 1201-1710, 1216-1451, 1217-1674, 1218-1674, 1222-1835, 1227-1761, 1227-1797, 1237-1356, 1237-1477, 1238-1759, 1238-1898, 1245-1477, 1248-1517, 1248-1771, 1256-1837, 1261-1488, 1266-1510, 1283-1817, 1285-1555, 1286-1814, 1293-1543, 1294-1997, 1297-2081, 1298-1589, 1321-1992, 1328-1601, 1355-1882, 1371-2013, 1381-1775, 1396-1734, 1410-1874, 1411-2059, 1411-2122,
	1413-2007, 1425-1795, 1425-1945, 1425-1949, 1425-2060, 1425-2070, 1425-2077, 1425-2110, 1425-2130, 1426-1604, 1427-2074, 1428-1824, 1428-1851, 1428-1889, 1428-1904, 1428-2031, 1428-2036, 1428-2081, 1428-2105, 1430-1778, 1431-1877, 1431-1890, 1431-2015, 1431-2025, 1431-2040, 1431-2064, 1431-2077, 1431-2106, 1440-1953, 1452-1676, 1457-1712, 1459-1725, 1460-2012, 1465-2011, 1466-2087, 1469-1673, 1469-1719, 1469-1931, 1471-1929, 1473-1659, 1474-2035, 1475-1673, 1475-1730, 1478-1980, 1489-1712, 1489-1723, 1489-1993, 1493-1747, 1496-1752, 1496-1993, 1497-1758, 1499-1719, 1499-2118, 1504-2118, 1505-2042, 1518-1768, 1518-1776, 1521-1894, 1523-1779, 1526-1741, 1535-2142, 1536-2089, 1539-1848, 1543-1766, 1549-1992, 1550-1836, 1552-1803, 1560-1790, 1560-1832, 1561-1861, 1564-2174, 1567-1715, 1571-1717, 1577-1803, 1579-2015, 1583-1871, 1584-2309, 1585-2067, 1587-1789, 1603-1845, 1603-1868, 1606-1938, 1606-1999, 1607-1778, 1615-1862, 1615-1880, 1616-2048, 1617-1890, 1619-2244, 1622-1861, 1623-2328, 1630-1853, 1630-1908, 1632-1837, 1632-2227, 1638-1884, 1643-1891, 1649-2018, 1649-2323, 1650-2324, 1660-1909, 1661-1881,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1664-1911, 1670-2416, 1676-1921, 1679-1872, 1679-1877, 1684-1954, 1688-1901, 1697-1980, 1699-2237, 1700-1949, 1705-1926, 1713-1944, 1717-2279, 1727-2381, 1729-2235, 1733-1956, 1735-1980, 1743-1984, 1743-1988, 1744-1966, 1744-1970, 1744-2355, 1746-2334, 1751-1978, 1752-2498, 1753-2028, 1753-2170, 1753-2212, 1763-2060, 1770-2047, 1770-2052, 1777-1989, 1779-2062, 1780-2058, 1786-2012, 1786-2033, 1789-2327, 1793-2094, 1797-2060, 1797-2063, 1801-2518, 1802-2513, 1811-2072, 1811-2103, 1814-2079, 1814-2510, 1816-2466, 1816-2494, 1820-2510, 1838-2087, 1844-2510, 1848-2382, 1850-2055, 1858-2143, 1859-2086, 1862-2049, 1863-2158, 1863-2510, 1863-2514, 1867-2510, 1867-2513, 1868-2510, 1870-2514, 1874-2513, 1874-2514, 1874-2533, 1877-2120, 1879-2129, 1879-2159, 1880-2112, 1880-2115, 1880-2122, 1880-2142, 1880-2148, 1882-2514, 1885-2165, 1885-2166, 1887-2490, 1888-2510, 1890-2127, 1900-2157, 1907-2117, 1907-2566, 1913-2514, 1915-2082, 1916-2165, 1917-2559, 1919-2139, 1924-2183, 1931-2199, 1933-2378, 1934-2219, 1935-2429, 1942-2140, 1943-2397, 1949-2388, 1951-2198, 1952-2230, 1958-2162, 1967-2237, 1969-2510, 1971-2196,
	1974-2623, 1975-2510, 1978-2185, 1978-2235, 1979-2219, 1979-2227, 1979-2261, 1981-2533, 1986-2563, 1987-2510, 1989-2579, 1991-2225, 1995-2662, 1997-2229, 2003-2262, 2005-2513, 2006-2426, 2007-2275, 2009-2514, 2015-2513, 2016-2477, 2021-2319, 2027-2289, 2033-2513, 2033-2539, 2033-2578, 2036-2275, 2038-2264, 2038-2279, 2038-2301, 2038-2308, 2038-2310, 2038-2327, 2039-2513, 2045-2510, 2048-2629, 2058-2326, 2061-2525, 2063-2316, 2066-2335, 2066-2816, 2078-2350, 2079-2699, 2080-2373, 2083-2646, 2086-2282, 2086-2321, 2086-2363, 2090-2404, 2090-2423, 2090-2513, 2091-2340, 2091-2349, 2091-2575, 2096-2336, 2098-2346, 2098-2350, 2098-2368, 2099-2294, 2100-2360, 2101-2339, 2101-2597, 2102-2790, 2103-2353, 2104-2335, 2104-2723, 2107-2406, 2110-2407, 2111-2626, 2112-2382, 2114-2626, 2115-2363, 2117-2569, 2122-2337, 2122-2362, 2124-2605, 2130-2374, 2133-2400, 2135-2480, 2139-2428, 2141-2461, 2145-2745, 2148-2410, 2156-2513, 2157-2375, 2159-2419, 2159-2554, 2159-2705, 2160-2414, 2167-2423, 2167-2444, 2167-2475, 2170-2344, 2171-2615, 2172-2487, 2172-2490, 2173-2392, 2173-2462, 2176-2359, 2176-2395, 2176-2460, 2180-2475,
	2182-2396, 2184-2466, 2191-2379, 2191-2501, 2194-2449, 2195-2552, 2204-2446, 2204-2456, 2205-2449, 2206-2456, 2215-2314, 2215-2471, 2216-2454, 2216-2485, 2220-2456, 2223-2479, 2226-2452, 2226-2461, 2226-2499, 2226-2799, 2226-2830, 2227-2430, 2227-2499, 2228-2437, 2228-2501, 2228-2521, 2230-2453, 2233-2435, 2233-2505, 2233-2604, 2234-2560, 2235-2456, 2235-2505, 2240-2803, 2243-2464, 2243-2476, 2243-2664, 2243-2665, 2249-2501, 2251-2501, 2253-2522, 2256-2441, 2256-2474, 2256-2555, 2259-2530, 2264-2474, 2267-2532, 2272-2668, 2282-2524, 2283-2544, 2283-2636, 2285-2519, 2293-2553, 2295-2493, 2295-2522, 2296-2515, 2296-2571, 2306-2571, 2306-2590, 2306-2839, 2307-2504, 2307-2569, 2309-2437, 2316-2577, 2316-2601, 2316-2607, 2323-2560, 2323-2587, 2331-2599, 2332-2591, 2332-2623, 2334-2580, 2334-2590, 2340-2604, 2342-2607, 2342-2615, 2343-2598, 2447-2896, 2483-3013, 2511-2717, 2527-2718, 2607-2978, 2647-2945, 2698-2945, 2699-2966, 2723-2888, 2764-2957

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
104/7510398CB1/ 1409	1-398, 1-407, 1-418, 1-552, 1-563, 1-568, 1-570, 1-574, 1-603, 1-625, 1-627, 1-797, 2-686, 2-755, 20-697, 20-837, 20-887, 22-45, 23-669, 23-675, 23-796, 23-797, 23-829, 23-838, 23-855, 23-869, 23-887, 23-889, 26-769, 198-337, 243-337, 303-329, 419-1090, 426-1090, 467-1090, 485-1090, 532-1350, 537-1350, 543-1090, 599-1409, 606-1090, 723-1214, 730-1214, 913-1214, 916-1214, 949-1214, 966-1214, 969-1214, 1064-1214, 1068-1214, 1325-1350, 1360-1383, 1361-1383, 1364-1383, 1384-1409
105/7510498CB1/ 4025	1-4025, 468-511, 567-610, 654-935, 841-1234, 842-1246, 842-1278, 891-1157, 1028-1572, 1028-1573, 1124-1495, 1162-1642, 1284-1773, 1401-2085, 1419-1654, 1474-2093, 1486-1794, 1504-2074, 1520-1857, 1522-2161, 1566-1982, 1582-1688, 1601-2402, 1603-2153, 1706-2233, 1707-2174, 1709-2003, 1709-2229, 1710-2218, 1712-2365, 1717-2329, 1722-1951, 1727-2014, 1731-2286, 1735-2024, 1756-2019, 1761-2026, 1770-2425, 1779-2236, 1780-2352, 1790-2043, 1790-2232, 1794-2071, 1799-2046, 1801-2028, 1801-2399, 1803-2263, 1806-2041, 1807-2263, 1808-2135, 1813-2451, 1824-2385, 1833-2105, 1837-1987, 1845-2101, 1845-2305, 1849-2101, 1854-2361, 1861-2388, 1861-2551, 1869-2377, 1875-2341, 1881-2521, 1884-1989, 1884-2366, 1884-2398, 1887-2048, 1887-2108, 1892-2446, 1900-2337, 1915-2469, 1916-2136, 1916-2400, 1922-2582, 1925-2194, 1927-2466, 1932-2220, 1934-2443, 1949-2184, 1950-2407, 1951-2407, 1955-2568, 1960-2494, 1960-2530, 1970-2089, 1970-2210, 1971-2492, 1971-2631, 1978-2210, 1981-2250, 1981-2504, 1989-2570, 1994-2221, 1999-2243, 2016-2550, 2018-2288, 2019-2547, 2026-2276, 2027-2730, 2030-2814, 2031-2322, 2054-2725, 2061-2334,
	2088-2615, 2104-2746, 2114-2508, 2129-2467, 2143-2607, 2144-2792, 2144-2855, 2146-2740, 2158-2528, 2158-2678, 2158-2682, 2158-2793, 2158-2803, 2158-2810, 2158-2843, 2158-2863, 2159-2337, 2160-2807, 2161-2557, 2161-2584, 2161-2622, 2161-2637, 2161-2764, 2161-2769, 2161-2814, 2161-2838, 2163-2511, 2164-2610, 2164-2623, 2164-2748, 2164-2758, 2164-2773, 2164-2797, 2164-2810, 2164-2839, 2173-2686, 2185-2409, 2190-2445, 2192-2458, 2193-2745, 2198-2744, 2199-2820, 2202-2406, 2202-2452, 2202-2664, 2204-2662, 2206-2392, 2207-2768, 2208-2406, 2208-2463, 2211-2713, 2222-2445, 2222-2456, 2222-2726, 2226-2480, 2229-2726, 2230-2491, 2232-2452, 2232-2851, 2237-2851, 2238-2775, 2251-2501, 2251-2509, 2254-2627, 2256-2512, 2259-2474, 2268-2875, 2269-2822, 2282-2725, 2285-2536, 2293-2523, 2293-2565, 2294-2594, 2297-2907, 2300-2448, 2304-2450, 2310-2536, 2312-2748, 2316-2604, 2317-3042, 2318-2800, 2320-2522, 2336-2578, 2336-2601, 2339-2671, 2339-2732, 2340-2511, 2348-2595, 2348-2613, 2349-2781, 2350-2623, 2352-2977, 2355-2594, 2356-3061, 2363-2586, 2363-2641, 2365-2570, 2365-2960, 2371-2617, 2376-2624, 2382-2751, 2382-3056, 2383-3057,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	2393-2642, 2394-2614, 2397-2644, 2403-3149, 2409-2654, 2412-2605, 2412-2610, 2417-2687, 2421-2634, 2430-2713, 2432-2970, 2433-2682, 2438-2659, 2446-2677, 2450-3012, 2460-3114, 2462-2968, 2466-2689, 2468-2713, 2476-2717, 2476-2721, 2477-2699, 2477-2703, 2477-3088, 2479-3067, 2484-2711, 2485-3231, 2486-2761, 2486-2903, 2486-2945, 2496-2793, 2503-2780, 2503-2785, 2510-2722, 2512-2795, 2513-2791, 2519-2745, 2519-2766, 2522-3060, 2526-2827, 2530-2793, 2530-2796, 2534-3251, 2535-3246, 2544-2805, 2544-2836, 2547-2812, 2547-3243, 2549-3199, 2549-3227, 2553-3243, 2571-2820, 2577-3243, 2581-3115, 2583-2788, 2591-2876, 2592-2819, 2595-2782, 2596-2891, 2596-3243, 2596-3247, 2600-3243, 2600-3246, 2601-3243, 2603-3247, 2607-3246, 2607-3247, 2607-3266, 2610-2853, 2612-2862, 2612-2892, 2613-2845, 2613-2848, 2613-2855, 2613-2875, 2613-2881, 2615-3247, 2618-2898, 2618-2899, 2620-3223, 2621-3243, 2623-2860, 2633-2890, 2640-2850, 2640-3299, 2646-3247, 2648-2815, 2649-2898, 2650-3292, 2652-2872, 2657-2916, 2664-2932, 2666-3111, 2667-2952, 2668-3162, 2675-2873, 2676-3130, 2682-3121, 2684-2931, 2685-2963, 2691-2895, 2700-2970,
	2702-3243, 2704-2929, 2707-3356, 2708-3243, 2711-2918, 2711-2968, 2712-2952, 2712-2960, 2712-2994, 2714-3266, 2719-3296, 2720-3243, 2722-3312, 2724-2958, 2728-3395, 2730-2962, 2736-2995, 2738-3246, 2739-3159, 2740-3008, 2742-3247, 2748-3246, 2749-3210, 2754-3052, 2760-3022, 2766-3246, 2766-3272, 2766-3311, 2769-3008, 2771-2997, 2771-3012, 2771-3034, 2771-3041, 2771-3043, 2771-3060, 2772-3246, 2778-3243, 2781-3362, 2791-3059, 2794-3258, 2796-3049, 2799-3068, 2799-3549, 2811-3083, 2812-3432, 2813-3106, 2816-3379, 2819-3015, 2819-3054, 2819-3096, 2823-3137, 2823-3156, 2823-3246, 2824-3073, 2824-3082, 2824-3308, 2825-3029, 2829-3069, 2831-3079, 2831-3083, 2831-3101, 2832-3027, 2833-3093, 2834-3072, 2834-3330, 2835-3523, 2836-3086, 2837-3068, 2837-3456, 2840-3139, 2843-3140, 2844-3359, 2845-3115, 2847-3359, 2848-3096, 2850-3302, 2855-3070, 2855-3095, 2857-3338, 2863-3107, 2866-3133, 2868-3213, 2872-3161, 2872-3186, 2874-3194, 2878-3134, 2878-3478, 2881-3143, 2889-3246, 2890-3108, 2892-3152, 2892-3287, 2892-3438, 2893-3147, 2900-3156, 2900-3177, 2900-3208, 2903-3077, 2904-3348, 2905-3220, 2905-3223, 2906-3125, 2906-3195,
	2909-3092, 2909-3128, 2909-3193, 2913-3208; 2915-3129, 2917-3199, 2922-3201, 2924-3112, 2924-3234, 2927-3182, 2928-3285, 2937-3179, 2937-3189, 2938-3182, 2939-3189, 2948-3047, 2948-3204, 2949-3187, 2949-3218, 2953-3189, 2956-3212, 2959-3185, 2959-3194, 2959-3232, 2959-3532, 2959-3563, 2960-3163, 2960-3232, 2961-3170, 2961-3234, 2961-3254, 2963-3186, 2966-3168, 2966-3238, 2966-3337, 2967-3293, 2968-3189, 2968-3238, 2973-3536, 2976-3197, 2976-3209, 2976-3397, 2976-3398, 2982-3234, 2984-3234, 2986-3255, 2989-3174, 2989-3207, 2989-3288, 2992-3263, 2997-3207, 3000-3265, 3005-3401, 3015-3257, 3016-3277, 3016-3369, 3018-3252, 3026-3286, 3028-3226, 3028-3255, 3029-3248, 3029-3304, 3039-3304, 3039-3323, 3039-3572, 3040-3237, 3040-3302, 3042-3170, 3049-3310, 3049-3334, 3049-3340, 3056-3293, 3056-3320, 3064-3332, 3065-3324, 3065-3356, 3067-3313, 3067-3323, 3073-3337, 3075-3340, 3075-3348, 3076-3331, 3180-3629, 3216-3746, 3307-3633, 3374-3616, 3380-3678, 3431-3678, 3432-3699, 3456-3621, 3487-3580, 3497-3690

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
106/7510044CB1/ 2951	1-525, 1-603, 1-2951, 219-642, 1864-2099, 2181-2609, 2629-2925, 2688-2951
107/7504509CB1/ 1832	1-206, 32-254, 33-275, 33-349, 37-246, 40-275, 42-1811, 44-275, 52-310, 52-357, 54-275, 56-360, 57-275, 79-317, 79-857, 79-877, 105-459, 105-554, 143-381, 155-402, 155-434, 155-650, 155-656, 157-611, 184-433, 203-580, 275-323, 275-486, 275-493, 275-508, 275-702, 275-736, 275-865, 293-799, 299-754, 309-799, 310-693, 311-780, 323-589, 333-889, 337-889, 349-599, 363-581, 363-639, 386-646, 390-894, 404-726, 405-938, 409-856, 410-762, 417-979, 422-684, 424-927, 427-700, 427-1110, 433-912, 433-993, 433-1050, 435-897, 439-1276, 447-596, 447-776, 449-852, 455-705, 458-737, 463-987, 466-972, 468-998, 471-1174, 485-771, 485-964, 497-1164, 499-1025, 500-987, 500-991, 501-726, 519-627, 520-627, 537-1034, 551-797, 551-814, 551-1028, 551-1047, 551-1048, 551-1076, 551-1085, 551-1107, 551-1160, 555-1188, 557-610, 559-1046, 561-1057, 573-1276, 587-1095, 589-855, 591-1112, 594-1094, 596-862, 599-889, 600-1086, 600-1138, 617-848, 618-1204, 628-1194, 634-1135, 634-1152, 635-901, 637-918, 638-1193, 641-840, 641-858, 641-918, 650-854,
	651-1389, 655-1170, 656-896, 657-925, 663-900, 671-953, 672-854, 684-1150, 685-1199, 686-1272, 686-1276, 689-900, 691-944, 691-1151, 691-1152, 691-1169, 691-1179, 691-1184, 691-1276, 691-1345, 691-1378, 692-1276, 695-1322; 697-1185, 700-863, 708-1169, 723-1440, 724-1307, 734-949, 748-1017, 754-1078, 766-1373, 768-1315, 773-984, 773-1267, 774-1013, 786-1021, 787-1292, 791-1415, 806-1275, 808-1323, 809-1169, 809-1331, 820-1187, 827-1093, 827-1094, 828-1276, 829-1094, 829-1130, 830-1327, 835-1114, 835-1276, 836-1415, 839-1276, 840-1275, 842-940, 842-1070, 842-1402, 845-1269, 851-1276, 851-1557, 856-1166, 856-1480, 856-1645, 857-1105, 857-1113, 860-1155, 861-1496, 867-1668, 869-1401, 869-1420, 871-1147, 873-1412, 874-1405, 878-1379, 878-1440, 883-1497, 884-1232, 884-1386, 884-1469, 887-1159, 887-1457, 888-1276, 888-1577, 889-1482, 889-1577, 890-1166, 891-1505, 892-1220, 907-1402, 912-1567, 918-1464, 931-1148, 931-1149, 933-1666, 935-1666, 936-1414, 937-1147, 939-1156, 940-1208, 940-1217,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	951-1528, 952-1282, 954-1552, 955-1208, 955-1457, 955-1555, 959-1557, 966-1614, 967-1219, 970-1260, 970-1811, 972-1276, 972-1315, 978-1624, 980-1474, 983-1258, 987-1249, 987-1575, 988-1276, 988-1585, 990-1276, 994-1276, 998-1224, 998-1267, 999-1208, 1001-1271, 1003-1545, 1004-1550, 1004-1806, 1006-1290, 1011-1407, 1013-1811, 1015-1526, 1019-1588, 1021-1811, 1022-1535, 1024-1811, 1025-1609, 1026-1288, 1027-1561, 1035-1669, 1037-1211, 1037-1510, 1040-1276, 1048-1650, 1050-1254, 1050-1287, 1053-1657, 1055-1341, 1055-1342, 1058-1430, 1058-1832, 1061-1811, 1064-1339, 1064-1363, 1064-1651, 1066-1666, 1067-1624, 1068-1811, 1071-1279, 1074-1251, 1077-1811, 1079-1370, 1079-1662, 1080-1658, 1092-1661, 1094-1377, 1095-1455, 1095-1620, 1098-1549, 1104-1652, 1105-1529, 1107-1811, 1110-1660, 1111-1400, 1112-1383, 1113-1626, 1113-1630, 1116-1670, 1117-1355, 1120-1377, 1122-1663, 1123-1669, 1125-1339, 1125-1627, 1128-1374, 1128-1637, 1128-1658, 1132-1500, 1136-1659, 1136-1666, 1137-1405, 1137-1670,
	1141-1670, 1150-1532, 1153-1549, 1157-1433, 1161-1669, 1164-1666, 1168-1625, 1169-1639, 1170-1571, 1173-1668, 1174-1414, 1174-1449, 1176-1668, 1183-1462, 1183-1670, 1184-1753, 1186-1549, 1186-1639, 1188-1349, 1189-1668, 1192-1665, 1194-1793, 1197-1479, 1197-1669, 1202-1409, 1205-1669, 1211-1375, 1211-1669, 1214-1652, 1216-1666, 1217-1535, 1217-1670, 1218-1641, 1218-1651, 1218-1823, 1219-1526, 1219-1665, 1220-1667, 1220-1669, 1222-1345, 1222-1468, 1223-1811, 1224-1665, 1225-1668, 1226-1667, 1226-1668, 1227-1665, 1227-1669, 1228-1379, 1228-1642, 1228-1668, 1228-1669, 1230-1666, 1230-1669, 1231-1364, 1231-1595, 1232-1435, 1232-1464, 1232-1668, 1233-1665, 1233-1824, 1234-1665, 1236-1669, 1236-1670, 1236-1803, 1237-1667, 1237-1669, 1244-1450, 1245-1666, 1247-1667, 1247-1669, 1249-1669, 1253-1385, 1253-1667, 1254-1669, 1254-1825, 1255-1527, 1255-1669, 1256-1667, 1256-1669, 1257-1669, 1258-1669, 1258-1798, 1265-1824, 1270-1654, 1271-1542, 1271-1653, 1280-1669, 1283-1669, 1288-1818, 1291-1669,
	1294-1669, 1295-1431, 1295-1669, 1296-1574, 1296-1652, 1296-1656, 1296-1666, 1296-1667, 1297-1536, 1299-1669, 1301-1542, 1301-1562, 1303-1667, 1303-1669, 1304-1598, 1307-1669, 1314-1638, 1315-1669, 1318-1669, 1322-1826, 1325-1613, 1325-1666, 1325-1826, 1326-1669, 1328-1569, 1329-1669, 1331-1615, 1331-1667, 1332-1665, 1333-1669, 1335-1735, 1337-1668, 1345-1552, 1345-1715, 1345-1821, 1347-1516, 1347-1825, 1348-1669, 1349-1811, 1350-1667, 1351-1801, 1353-1668, 1356-1664, 1359-1825, 1362-1625, 1362-1668, 1363-1633, 1364-1661, 1364-1668, 1364-1690, 1365-1605, 1368-1810, 1372-1669, 1373-1640, 1377-1770, 1379-1667, 1383-1643, 1385-1667, 1387-1643, 1387-1668, 1389-1668, 1389-1825, 1404-1808, 1408-1479, 1408-1765, 1414-1668, 1416-1665, 1425-1668, 1425-1669, 1439-1666, 1440-1689, 1444-1669, 1447-1667, 1454-1810, 1457-1668, 1470-1667, 1478-1666, 1478-1669, 1481-1665, 1484-1666, 1485-1810, 1509-1666, 1516-1810, 1518-1810, 1525-1809, 1528-1667, 1528-1825, 1533-1810, 1535-1634, 1539-1668, 1540-1810, 1545-1811, 1547-1810, 1560-1824, 1573-1807, 1586-1810, 1600-1667, 1624-1773, 1624-1810, 1695-1832

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
108/7506825CB1/ 1471	1-240, 1-1462, 23-254, 29-286, 52-260, 52-357, 78-900, 80-228, 106-684, 268-707, 285-982, 286-534, 287-1006, 295-913, 301-579, 306-825, 320-612, 335-555, 344-955, 349-979, 355-885, 363-792, 379-756, 380-902, 391-1078, 397-642, 398-942, 403-1391, 405-681, 408-1391, 411-1253, 416-1031, 418-672, 421-1078, 422-1391, 439-690, 441-1023, 447-965, 450-942, 460-751, 471-778, 473-737, 474-845, 475-1391, 476-1391, 480-713, 481-1391, 488-721, 497-813, 499-731, 499-757, 504-1201, 511-1391, 514-1243, 522-629, 522-739, 528-1391, 534-755, 538-988, 544-736, 545-1150, 549-1150, 550-1202, 557-1411, 561-726, 567-973, 569-1055, 582-1449, 591-1051, 592-944, 599-1172, 603-847, 623-1163, 625-1304, 626-736, 631-1069, 632-960, 632-1417, 650-1216, 659-1192, 664-1036, 664-1321, 666-1177, 668-933, 668-943, 668-1209, 673-1346, 675-915, 675-922, 678-1447, 685-995, 687-1341, 689-1376, 690-920, 690-1230, 696-1200, 697-951, 697-980, 708-1391, 711-1353, 713-1362, 715-1167, 715-1267, 716-1344, 717-1391, 721-1248, 726-981, 727-1442,
	735-984, 743-1425, 744-1188, 752-1463, 757-1021, 760-1002, 762-1307, 762-1378, 762-1395, 772-1058, 772-1391, 781-1057, 781-1460, 784-1364, 789-986, 790-1013, 798-1149, 801-1200, 804-1421, 805-1461, 811-1465, 814-1054, 814-1060, 819-999, 826-1419, 830-1093, 838-1051, 846-1088, 846-1273, 846-1380, 847-1128, 857-1461, 860-1429, 864-1437, 869-1107, 870-1416, 872-1461, 883-1206, 883-1427, 906-1357, 911-1471, 913-1354, 916-1157, 925-1460, 928-1205, 931-1396, 935-1183, 936-1207, 936-1212, 951-1456, 952-1465, 955-1100, 958-1071, 960-1465, 963-1238, 964-1449, 970-1224, 993-1252, 996-1469, 1001-1195, 1013-1464, 1017-1427, 1020-1439, 1020-1448, 1025-1465, 1030-1214, 1030-1462, 1031-1465, 1032-1448, 1034-1462, 1034-1471, 1038-1447, 1046-1450, 1053-1447, 1055-1295, 1058-1447, 1060-1391, 1060-1436, 1064-1319, 1067-1465, 1082-1268, 1090-1449, 1102-1332, 1102-1445, 1112-1444, 1117-1445, 1123-1418, 1125-1382, 1128-1445, 1134-1427, 1148-1445, 1150-1447, 1151-1425, 1160-1445, 1166-1447, 1171-1390, 1173-1451, 1173-1469, 1174-1447, 1178-1425, 1180-1442, 1186-1380, 1188-1451, 1201-1382, 1209-1441, 1213-1449, 1258-1447, 1260-1437, 1308-1444, 1338-1449
109/7506828CB1/ 1344	1-240, 1-1332, 23-254, 23-298, 24-577, 29-260, 31-311, 33-536, 52-276, 52-307, 52-310, 59-362, 59-396, 59-397, 59-584, 61-758, 63-541, 65-355, 71-704, 72-305, 72-319, 72-326, 72-353, 72-505, 72-528, 72-531, 72-545, 72-572, 72-573, 72-579, 72-588, 72-590, 72-599, 72-609, 72-658, 72-698, 72-710, 72-713, 72-724, 72-761, 73-328, 73-349, 78-815, 78-905, 79-354, 79-557, 79-570, 80-228, 80-321, 80-325, 80-339, 80-343, 80-345, 80-360, 80-377, 80-383, 84-361, 86-586, 88-542, 90-664, 95-352, 96-361, 96-363, 96-366, 96-380, 96-420, 98-382, 98-386, 99-326, 100-374, 100-394, 100-668, 102-389, 104-361, 104-532, 105-337, 105-354, 105-378, 105-385, 105-417, 105-418, 106-360, 106-376, 106-379, 106-387, 106-400, 106-412, 106-579, 106-704, 107-258, 107-331, 107-337, 107-342, 107-370, 107-392, 107-397, 107-411, 107-440, 107-623, 107-687, 107-696, 107-702, 107-715, 107-724, 107-800, 108-395, 110-403, 110-664, 110-717, 110-727, 110-738, 111-433, 111-543, 111-614, 111-673, 111-696, 111-751, 112-517, 116-384, 116-725, 117-727, 118-361, 118-498,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	121-364, 121-634, 121-665, 123-679, 124-471, 127-667, 127-677, 128-393, 129-567, 132-709, 145-666, 149-726, 171-808, 173-308, 180-780, 183-466, 186-544, 186-610, 222-514, 227-555, 258-391, 266-642, 267-434, 276-542, 280-481, 333-772, 351-599, 364-610, 366-644, 384-642, 385-677, 388-654, 399-690, 438-620, 440-733, 444-821, 445-711, 445-821, 462-707, 470-746, 483-737, 504-755, 517-769, 525-816, 536-822, 538-802, 545-778, 553-786, 553-789, 564-796, 564-822, 587-694, 587-804, 599-820, 609-801, 612-1262, 626-791, 691-801, 744-1317, 822-1053, 822-1077, 822-1082, 822-1266, 822-1319, 822-1341, 825-970, 826-1286, 828-941, 830-1342, 833-1108, 834-1338, 840-1094, 842-1329, 848-1335, 849-1331, 860-1337, 863-1122, 863-1321, 866-1287, 866-1316, 868-1316, 871-1065, 877-1296, 883-1334, 883-1343, 888-1297, 890-1309, 890-1318, 895-1337, 900-1084, 900-1332, 901-1340, 902-1318, 904-1335, 904-1337, 904-1344, 908-1316, 908-1317, 916-1319, 923-1317, 925-1165, 928-1317, 934-1189, 937-1341, 951-1110, 952-1138, 953-1297, 960-1323, 972-1202, 972-1315, 982-1321, 983-1261, 987-1315, 993-1288, 995-1252, 998-1315, 1003-1341, 1004-1297, 1018-1315,
	1020-1317, 1021-1295, 1021-1315, 1030-1315, 1036-1317, 1041-1260, 1043-1319, 1043-1337, 1044-1317, 1048-1295, 1050-1312, 1056-1250, 1058-1319, 1071-1252, 1079-1311, 1082-1306, 1083-1322, 1084-1314, 1090-1297, 1092-1314, 1095-1319, 1102-1319, 1119-1311, 1128-1317, 1130-1307, 1152-1321, 1159-1319, 1161-1316, 1165-1320, 1178-1314, 1208-1321, 1237-1294
110/7506831CB1/ 1429	1-240, 1-1422, 23-254, 29-260, 52-260, 78-756, 78-960, 78-961, 80-228, 100-260, 107-258, 107-260, 191-459, 259-494, 259-505, 259-873, 259-942, 259-966, 261-539, 266-785, 279-537, 280-572, 283-549, 294-585, 295-515, 304-915, 309-939, 315-845, 323-752, 331-824, 335-628, 338-768, 339-716, 339-768, 340-606, 340-635, 357-602, 358-902, 365-641, 376-991, 378-632, 381-1038, 396-1069, 399-650, 401-983, 407-925, 410-902, 412-664, 420-711, 431-738, 433-697, 434-805, 435-1351, 436-1351, 440-673, 441-1351, 448-681, 448-684, 457-773, 459-691, 471-1351, 474-1038, 474-1203, 482-589, 482-699, 488-1351, 494-715, 498-948, 504-696, 505-1110, 510-1162, 517-1371, 518-717, 521-686, 527-933, 529-1015, 551-1011, 552-904, 559-1132, 563-807, 577-1309, 583-1123, 585-1264, 586-696, 591-1029, 592-920, 592-1377, 594-1161, 600-1351, 601-1110, 610-1176, 619-1152, 624-996, 624-1281, 626-1137, 628-893, 628-903, 628-1169, 631-1351, 633-1306, 635-875, 635-882, 638-1426, 645-955, 647-1301, 649-1336, 650-880, 650-1034, 656-945, 657-911,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	657-940, 668-1351, 671-1313, 673-1322, 675-1127, 675-1227, 676-1304, 677-1351, 678-1351, 681-1208, 686-941, 687-1402, 695-944, 703-1322, 704-1148, 712-1429, 717-981, 720-962, 722-1267, 722-1338, 722-1355, 732-1018, 732-1351, 741-1017, 741-1420, 744-1324, 749-946, 750-973, 758-1109, 761-1160, 764-1381, 765-1421, 771-1425, 774-1014, 774-1020, 779-959, 786-1379, 790-1053, 798-1011, 806-1048, 806-1233, 806-1340, 807-1088, 817-1421, 820-1389, 824-1397, 829-1067, 832-1421, 843-1166, 843-1387, 866-1317, 873-1314, 876-1117, 885-1420, 891-1356, 895-1143, 896-1165, 896-1167, 896-1172, 911-1409, 912-1429, 915-1060, 918-1031, 920-1429, 923-1198, 924-1428, 928-1376, 930-1184, 953-1212, 956-1377, 961-1155, 973-1424, 980-1408, 985-1427, 990-1174, 991-1429, 998-1407, 1006-1409, 1013-1407, 1018-1407, 1024-1279, 1027-1429, 1029-1351, 1041-1200, 1042-1228, 1050-1413, 1072-1411, 1083-1378, 1092-1396, 1094-1387, 1111-1385, 1111-1405, 1120-1405, 1126-1407, 1131-1350, 1133-1409, 1133-1427, 1165-1427, 1169-1401, 1173-1412, 1182-1404, 1192-1409, 1249-1409, 1251-1406, 1268-1404, 1327-1384
111/7509968CB1/ 2187	1-252, 1-265, 1-517, 1-563, 1-1959, 79-333, 204-335, 257-952, 300-574, 470-998, 521-998, 657-1143, 657-1287, 665-998, 668-1003, 669-1199, 681-1083, 685-1362, 685-1434, 686-1239, 693-1300, 693-1335, 696-1256, 697-1313, 706-994, 707-1246, 708-1271, 709-1239, 734-1384, 735-1246, 744-1393, 751-1094, 763-1386, 776-1368, 789-1335, 791-1063, 793-1491, 799-1426, 799-1427, 802-1435, 805-1378, 807-1428, 810-1204, 817-1259, 823-1447, 824-1388, 824-1393, 826-1445, 836-922, 841-1141, 849-1453, 856-1276, 857-1474, 858-1329, 859-1444, 859-1488, 870-1459, 875-1334, 881-1278, 883-1255, 884-1486, 895-1452, 903-1560, 904-1512, 925-1082, 925-1222, 925-1573, 926-1308, 926-1455, 933-1222, 933-1440, 941-1459, 942-1386, 944-1459, 948-1539, 960-1532, 962-1518, 969-1386, 975-1599, 983-1622, 997-1530, 1017-1625, 1023-1671, 1040-1222, 1045-1416, 1052-1599, 1054-1470, 1060-1309, 1061-1325, 1061-1593, 1066-1621, 1106-1414, 1106-1495, 1106-1648, 1106-1652, 1106-1694, 1127-1779, 1143-1345, 1148-1335, 1181-1715, 1193-1722, 1205-1746, 1216-1736, 1263-1709, 1288-1880, 1327-1812, 1334-1878, 1342-1608, 1376-1621, 1527-1747, 1545-2187, 1586-2034, 1608-1880, 1668-1943, 1750-1944
112/7510232CB1/ 2020	1-2020, 898-1402, 901-1176, 901-1237, 901-1410, 902-1356, 903-1110, 904-1110, 910-1146, 928-1532, 1247-1438, 1247-1440, 1250-1437, 1251-1440, 1274-1293, 1297-1883, 1605-1784, 1949-2020
113/7510233CB1/ 2167	1-1927, 858-1574, 885-1313, 948-1156, 948-1439, 949-1201, 949-1272, 1012-1630, 1051-1562, 1125-1718, 1126-1632, 1154-1637, 1154-1671, 1156-1347, 1156-1349, 1159-1346, 1159-1560, 1159-1671, 1160-1349, 1170-1486, 1181-1549, 1183-1202, 1191-1671, 1197-1503, 1206-1792, 1209-1763, 1217-1651, 1225-2030, 1229-1671, 1232-1794, 1277-1779, 1278-1523, 1289-1671, 1295-1833, 1335-1801, 1344-1864, 1377-1997, 1380-1800, 1382-1671, 1431-1782, 1443-1976, 1443-1977, 1452-1573, 1473-1700, 1488-1767, 1495-1737, 1503-1738, 1507-1997, 1514-1693, 1542-1976, 1542-2025, 1545-1812, 1553-1800, 1563-1747, 1563-2070, 1618-2167, 1621-1815, 1725-2016, 1729-1970, 1737-2004

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
114/7510304CB1/ 1706	1-238, 1-1704, 101-326, 112-367, 123-443, 124-329, 369-604, 369-685, 369-692, 369-885, 391-697, 411-870, 460-725, 473-1006, 474-1086, 488-1064, 488-1082, 493-756, 504-711, 504-1087, 519-1093, 522-810, 534-922, 572-1175, 588-1161, 606-706, 609-871, 613-743, 618-1040, 621-1153, 700-1273, 713-1320, 717-925, 723-1680, 728-1055, 728-1170, 728-1203, 728-1240, 728-1276, 743-1025, 767-1302, 782-1681, 788-1679, 796-1039, 803-1286, 807-1090, 808-1342, 811-1064, 815-1080, 819-1159, 821-1395, 822-1094, 822-1301, 824-974, 831-1676, 832-1443, 846-1428, 850-1107, 866-984, 889-1343, 892-1040, 908-1171, 910-1399, 914-1239, 914-1460, 914-1465, 914-1511, 914-1566, 923-1502, 924-1201, 936-1680, 941-1680, 945-1163, 949-1198, 949-1302, 958-1668, 962-1680, 973-1291, 973-1498, 975-1344, 975-1365, 985-1339, 993-1112, 995-1214, 1005-1265, 1008-1162, 1009-1703, 1019-1524, 1021-1495, 1022-1280, 1022-1287, 1024-1197, 1034-1459, 1044-1469, 1049-1676, 1066-1336, 1081-1388, 1084-1396, 1110-1192, 1110-1706, 1116-1705, 1119-1656, 1121-1371, 1126-1525, 1130-1705, 1134-1375, 1140-1243, 1140-1426, 1141-1372, 1145-1706, 1146-1350, 1150-1410, 1152-1392,
	1154-1682, 1164-1442, 1170-1451, 1195-1562, 1197-1706, 1213-1599, 1224-1423, 1234-1700, 1242-1467, 1244-1694, 1247-1692, 1247-1703, 1256-1524, 1259-1697, 1259-1706, 1265-1694, 1265-1706, 1267-1696, 1273-1449, 1274-1619, 1274-1706, 1275-1706, 1276-1484, 1276-1694, 1276-1695, 1277-1697, 1280-1706, 1282-1697, 1288-1696, 1289-1706, 1295-1475, 1299-1686, 1308-1559, 1310-1702, 1311-1701, 1322-1466, 1327-1598, 1327-1694, 1337-1702, 1338-1416, 1363-1697, 1381-1701, 1386-1686, 1399-1700, 1400-1694, 1407-1640, 1412-1497, 1427-1606, 1428-1694, 1451-1654, 1478-1698, 1480-1706, 1485-1679, 1493-1694, 1504-1654, 1512-1706, 1520-1706, 1527-1694, 1531-1704, 1535-1623, 1548-1706, 1556-1694, 1565-1691, 1566-1706, 1580-1694, 1644-1706
115/7510461CB1/ 2742	1-299, 1-447, 1-474, 1-515, 1-620, 1-720, 6-238, 6-474, 7-895, 7-2478, 8-430, 12-517, 17-299, 17-726, 23-473, 23-474, 23-686, 24-275, 25-631, 63-307, 78-890, 168-474, 168-800, 168-1037, 174-1002, 181-822, 186-857, 186-863, 192-830, 207-775, 216-799, 220-536, 224-808, 230-818, 232-794, 233-874, 241-841, 249-899, 259-474, 266-469, 273-877, 273-902, 282-830, 286-1021, 299-546, 299-598, 306-578, 312-1022, 323-557, 324-474, 332-937, 340-1020, 348-907, 365-965, 366-1019, 366-1023, 367-987, 375-1019, 383-986, 389-895, 399-474, 403-474, 430-1022, 439-1100, 443-747, 452-1069, 454-1129, 462-967, 474-1030, 478-1127, 482-921, 485-1048, 486-869, 489-1087, 509-1062, 518-742, 519-1112, 523-1147, 526-1143, 527-1143, 531-849, 542-1069, 549-1030, 558-829, 573-1017, 573-1021, 575-1366, 600-832, 608-886, 610-1441, 620-877, 621-891, 633-903, 641-1012, 642-882, 642-928, 661-1363, 666-965, 668-1188, 675-873, 675-1012, 693-1021, 709-1599, 717-989, 740-1455, 752-1684, 755-1385, 823-1102, 934-1419, 937-1021, 939-1021, 939-1022,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	958-1021, 990-1102, 1055-1147, 1078-1599, 1078-1607, 1078-1752, 1168-1897, 1247-1677, 1247-1964, 1247-1968, 1248-1982, 1282-2095, 1422-2196, 1425-1936, 1430-2109, 1441-2242, 1517-2196, 1518-2145, 1530-2160, 1530-2200, 1531-2329, 1550-2454, 1558-2116, 1613-2404, 1613-2421, 1621-2396, 1628-2411, 1640-2284, 1640-2420, 1647-2010, 1656-2374, 1694-2455, 1700-2342, 1719-2455, 1722-2001, 1747-1898, 1771-2038, 1783-2419, 1784-2039, 1794-2424, 1851-2520, 1903-2191, 1909-2430, 1909-2492, 1935-2192, 1936-2493, 1946-2435, 1956-2213, 1957-2204, 1960-2381, 1978-2198, 2006-2494, 2025-2483, 2043-2444, 2047-2365, 2068-2476, 2069-2478, 2075-2316, 2091-2333, 2102-2304, 2102-2481, 2107-2742, 2113-2239, 2120-2385, 2121-2488, 2152-2403, 2156-2478, 2162-2445, 2165-2497, 2166-2431, 2171-2478, 2189-2478, 2194-2396, 2194-2442, 2194-2478, 2208-2463, 2213-2489, 2251-2538, 2257-2498
116/7510392CB1/ 1362	1-261, 1-278, 1-480, 1-492, 6-255, 6-267, 6-301, 6-412, 6-492, 6-1360, 8-564, 210-456, 220-494, 222-497, 233-372, 233-492, 233-500, 238-525, 239-485, 239-488, 239-490, 239-522, 239-545, 241-342, 241-434, 241-493, 241-503, 241-523, 241-564, 243-471, 243-475, 243-484, 243-487, 243-492, 243-495, 243-503, 243-509, 243-516, 243-540, 245-492, 245-512, 253-380, 253-541, 253-559, 258-490, 258-499, 261-496, 261-506, 261-512, 261-528, 263-500, 264-540, 265-494, 265-534, 266-513, 274-482, 280-424, 286-406, 286-520, 289-416, 289-419, 297-503, 297-532, 309-499, 309-507, 309-531, 313-525, 475-748, 506-766, 687-935, 706-1259, 710-1138, 715-1320, 719-1293, 767-1360, 772-993, 781-1360, 787-1054, 797-1360, 798-1033, 806-1356, 818-1138, 819-1096, 820-1063, 821-1278, 824-1049, 825-1338, 827-1362, 832-1343, 837-1360, 845-1358, 849-1043, 849-1056, 849-1082, 849-1360, 855-1278, 859-1124, 860-1115, 860-1121, 863-1360, 866-1360, 869-1359, 873-1115, 875-1360, 876-1132, 876-1208, 883-1145, 888-1360, 889-1303, 889-1360, 901-1360,
	915-1360, 918-1178, 919-1324, 919-1353, 919-1360, 920-1360, 922-1360, 923-1191, 925-1360, 926-1356, 934-1360, 937-1360, 942-1360, 948-1360, 950-1360, 951-1360, 952-1360, 956-1360, 965-1360, 966-1360, 984-1360, 988-1360, 990-1360, 1004-1360, 1007-1171, 1007-1340, 1007-1360, 1009-1360, 1025-1360, 1026-1242, 1035-1360, 1039-1360, 1044-1360, 1048-1360, 1050-1360, 1057-1360, 1071-1323, 1071-1360, 1074-1360, 1077-1299, 1084-1360, 1101-1360, 1108-1360, 1112-1360, 1113-1360, 1118-1360, 1126-1278, 1126-1360, 1130-1360, 1135-1360, 1140-1256, 1144-1360, 1150-1360, 1153-1344, 1153-1360, 1156-1360, 1162-1360, 1216-1361, 1226-1360, 1227-1360, 1229-1360, 1231-1360, 1251-1360, 1262-1360, 1265-1360, 1269-1360

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
59	7313196CB1	CARCTXT01
60	6465289CB1	BRABDIR01
61	7506357CB1	ADRETUE02
62	6878857CB1	PANCNOT04
63	7506021CB1	BRAGNON02
64	7503356CB1	BRAINOR03
65	7509052CB1	UCMCL5T01
66	7503366CB1	EPIMNON05
67	7505933CB1	TLYMNOT08
69	1439986CB1	THYRNOT03
70	2008979CB1	OVARTUT02
72	7506782CB1	BRAINON01
73	7506941CB1	NEUTFMT01
74	7507072CB1	OVARTUT05
76	7509097CB1	TONSDIT01
77	7509118CB1	TONSDIT01
78	7509312CB1	TONSDIT01
79	90126902CB1	NERDTDN03
80	7509352CB1	NPOLNOT01
82	7509367CB1	PITUNON01
83	7500455CB1	SKINDIA01
84	7510401CB1	KIDNNOT19
85	7504702CB1	LUNGAST01
86	7509113CB1	LUNGFET03
87	7509140CB1	BLADTUT07
88	7509223CB1	LUNGFET03
89	7509272CB1	TONSDIT01
90	7509327CB1	TONSDIT01
91	7504677CB1	BONEUNT01
92	7504534CB1	SCOMDIC01

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
93	7507771CB1	MCLDTXN05
94	7504732CB1	BRABNOE02
95	950917CB1	HEAANOT01
96	7459720CB1	PROSNOT11
97	7503300CB1	PROSTUS23
98	7503334CB1	BRSTNOT02
99	7503341CB1	LUNGTUT08
100	7509936CB1	KIDEUNE02
101	7509986CB1	PENITUT01
102	7510010CB1	BRSTTUT01
103	7510056CB1	PROTDNV06
105	7510498CB1	PROTDNV06
106	7510044CB1	PLACFEB01
107	7504509CB1	COLCTUT02
108	7506825CB1	SINTNOR01
109	7506828CB1	KIDNNOT26
110	7506831CB1	SINTNOR01
111	7509968CB1	ESOGTME01
112	7510232CB1	MCLDTXN05
113	7510233CB1	SYNORAB01
114	7510304CB1	BRAHTDK01
115	7510461CB1	HIPONOT01
116	7510392CB1	BRSTNOT02

Table 6

Library	Vector	Library Description
ADRETUE02	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from right adrenal tumor tissue removed from a 49-year-old Caucasian male during unilateral adrenalectomy. Pathology indicated adrenal cortical carcinoma comprising nearly the entire specimen. The tumor was attached to the adrenal gland which showed mild cortical atrophy. The tumor was encapsulated, being surrounded by a thin (1-3 mm) rim of connective tissue. The patient presented with adrenal cancer, abdominal pain, pyrexia of unknown origin, and deficiency anemia. Patient history included benign hypertension. Previous surgeries included adenotonsillectomy. Patient medications included aspirin, calcium, and iron. Family history included atherosclerotic coronary artery disease in the mother; cerebrovascular accident and atherosclerotic coronary artery disease in the father; and benign hypertension in the grandparent(s).
BLADTUT07	pINCY	Library was constructed using RNA isolated from bladder tumor tissue removed from the anterior bladder wall of a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrostomy. Pathology indicated a grade 3 transitional cell carcinoma in the left lateral bladder. Patient history included angina, emphysema, and tobacco use. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
BONEUNT01	pINCY	Library was constructed using RNA isolated from Saos-2, a primary osteogenic sarcoma cell line (ATCC HTB-85) derived from an 11-year-old Caucasian female.
BRABDIR01	pINCY	Library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Patient history included Huntington's disease, emphysema, and tobacco abuse.
BRABNOE02	PBK-CMV	This 5' biased random primed library was constructed using RNA isolated from vermis tissue removed from a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly, and an enlarged spleen and liver. Patient medications included simethicone, Lasix, Digoxin, Colace, Zantac, captopril, and Vasotec.

Table 6

Library	Vector	Library Description
BRAGNON02	pINCY	This normalized substantia nigra tissue library was constructed from 4.2 10e7 independent clones from a substantia nigra tissue library. Starting RNA was made from RNA isolated from substantia nigra tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were pairs of Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles.
		The posterior hippocampus contained a microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloidal goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAHTDK01	PSPORT1	This amplified and normalized library was constructed using pooled RNA isolated from archaecortex, anterior and posterior hippocampus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver. 7.6x10e5 independent clones from this amplified library were normalized in 1 round using conditions adapted Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.

Table 6

Library	Vector	Library Description
BRAINON01	PSPORT1	Library was constructed and normalized from 4.88 million independent clones from a brain tissue library. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right fronto-parietal part of the brain. The normalization and hybridization conditions were adapted from Soares et al., PNAS (1994) 91:9228, except that a significantly longer (48-hour) reannealing hybridization was used.
BRAINOR03	PBK-CMV	This random primed library was constructed using pooled cDNA from two donors. cDNA was generated using mRNA isolated from brain tissue removed from a Caucasian male fetus (donor A) who was stillborn with a hypoplastic left heart at 23 weeks' gestation and from brain tissue removed from a Caucasian male fetus (donor B), who died at 23 weeks' gestation from premature birth. Serologies were negative for both donors and family history for donor B included diabetes in the mother.
BRSTNOT02	PSPORT1	Library was constructed using RNA isolated from diseased breast tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated proliferative fibrocysytic changes characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Pathology for the associated tumor tissue indicated an invasive grade 4 mammary adenocarcinoma. Patient history included atrial tachycardia and a benign neoplasm. Family history included cardiovascular and cerebrovascular disease.
BRSTTUT01	PSPORT1	Library was constructed using RNA isolated from breast tumor tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma of mixed lobular and ductal type, extensively involving the left breast. The tumor was identified in the deep dermis near the lactiferous ducts with extracapsular extension. Seven mid and low and five high axillary lymph nodes were positive for tumor. Proliferative fibrocysytic changes were characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Patient history included atrial tachycardia, blood in the stool, and a benign breast neoplasm. Family history included benign hypertension, atherosclerotic coronary artery disease, cerebrovascular disease, and depressive disorder.
CARCTXT01	PSPORT1	Library was constructed using RNA from chondrocytes that were isolated from pooled knee cartilage obtained during total knee joint replacement. The cartilage was removed from the underlying bone, chopped into smaller pieces, and stimulated with 5 ng/ml IL-1 for 18 hours.

Table 6

Library	Vector	Library Description
COLCTUT02	pINCY	Library was constructed using RNA isolated from colon tumor tissue removed from the cecum of a 30-year-old Caucasian female during partial colectomy, open liver biopsy, incidental appendectomy, and permanent colostomy. Pathology indicated carcinoid tumor (grade 1 neuroendocrine carcinoma) arising in the terminal ileum, forming a mass in the right colon. Patient history included chronic sinus infections and endometriosis. Family history included hyperlipidemia, anxiety, upper lobe lung cancer, stomach cancer, liver cancer, and cirrhosis.
EPIMNON05	pINCY	This normalized mammary epithelial cell tissue library was constructed from 3.28 million independent clones from an epithelial cell tissue library. Starting RNA was made from untreated mammary epithelial cell tissue removed from a 21-year-old female. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 -hours/round) reannealing hybridization was used.
ESOGTME01	PSPORT	This 5' biased random primed library was constructed using RNA isolated from esophageal tissue removed from a 53-year-old Caucasian male during a partial esophagectomy, proximal gastrectomy, and regional lymph node biopsy. Pathology indicated no significant abnormality in the non-neoplastic esophagus. Pathology for the matched tumor tissue indicated invasive grade 4 (of 4) adenocarcinoma, forming a sessile mass situated in the lower esophagus, 2 cm from the gastroesophageal junction and 7 cm from the proximal margin. The tumor invaded through the muscularis propria into the adventitial soft tissue. Metastatic carcinoma was identified in 2 of 5 paragastric lymph nodes with perinodal extension. The patient presented with dysphagia. Patient history included membranous nephritis, hyperlipidemia, benign hypertension, and anxiety state. Previous surgeries included an adenotonsillectomy, appendectomy, and inguinal hernia repair. The patient was not taking any medications.
		Family history included atherosclerotic coronary artery disease, alcoholic cirrhosis, alcohol abuse, and an abdominal aortic aneurysm rupture in the father; breast cancer in the mother; a myocardial infarction and atherosclerotic coronary artery disease in the sibling(s); and myocardial infarction and atherosclerotic coronary artery disease in the grandparent(s).

Table 6

Library	Vector	Library Description
HEAANOT01	pINCY	Library was constructed using RNA isolated from right coronary and right circumflex coronary artery tissue removed from the explanted heart of a 46-year-old Caucasian male during a heart transplantation. Patient history included myocardial infarction from total occlusion of the left anterior descending coronary artery, atherosclerotic coronary artery disease, hyperlipidemia, myocardial ischemia, dilated cardiomyopathy, left ventricular dysfunction, and tobacco abuse. Previous surgeries included cardiac catheterization. Family history included atherosclerotic coronary artery disease.
HIPONOT01	PBLUESCRIPT	Library was constructed using RNA isolated from the hippocampus tissue of a 72-year-old Caucasian female who died from an intracranial bleed. Patient history included nose cancer, hypertension, and arthritis.
KIDEUNE02	pINCY	This 5' biased random primed library was constructed using RNA isolated from an untreated transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue (Invitrogen). The cells were transformed with adenovirus 5 DNA.
KIDNNOT19	pINCY	Library was constructed using RNA isolated from kidney tissue of a 65-year-old Caucasian male during an exploratory laparotomy and nephroureterectomy. Pathology for the associated tumor tissue indicated a grade 1 renal cell carcinoma within the upper pole of the left kidney. Patient history included malignant melanoma of the abdominal skin, benign neoplasm of the colon, cerebrovascular disease, and umbilical hernia. Family history included myocardial infarction, atherosclerotic coronary artery disease, cerebrovascular disease, prostate cancer, myocardial infarction, and atherosclerotic coronary artery disease.
KIDNNOT26	pINCY	Library was constructed using RNA isolated from left kidney medulla and cortex tissue removed from a 53-year-old Caucasian female during a nephroureterectomy. Pathology for the associated tumor tissue indicated grade 2 renal cell carcinoma involving the lower pole of the kidney. Patient history included hyperlipidemia, cardiac dysrhythmia, metrorrhagia, normal delivery, cerebrovascular disease, atherosclerotic coronary artery disease, and tobacco abuse. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
LUNGAST01	PSPORT1	Library was constructed using RNA isolated from the lung tissue of a 17-year-old Caucasian male, who died from head trauma. Patient history included asthma.
LUNGFET03	pINCY	Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.

Table 6

Library	Vector	Library Description
LUNGTUT08	pINCY	Library was constructed using RNA isolated from lung tumor tissue removed from a 63-year-old Caucasian male during a right upper lobectomy with fiberoptic bronchoscopy. Pathology indicated a grade 3 adenocarcinoma. Patient history included atherosclerotic coronary artery disease, an acute myocardial infarction, rectal cancer, an asymtomatic abdominal aortic aneurysm, tobacco abuse, and cardiac dysrhythmia. Family history included congestive heart failure, stomach cancer, and lung cancer, type II diabetes, atherosclerotic coronary artery disease, and an acute myocardial infarction.
MCLDTXN05	pINCY	This normalized dendritic cell library was constructed from 1 million independent clones from a pool of two derived dendritic cell libraries. Starting libraries were constructed using RNA isolated from untreated and treated derived dendritic cells from umbilical cord blood CD34+ precursor cells removed from a male. The cells were derived with granulocyte/macrophage colony stimulating factor (GM-CSF), tumor necrosis factor alpha (TNF alpha), and stem cell factor (SCF). The GM-CSF was added at time 0 at 100 ng/ml, the TNF alpha was added at time 0 at 2.5 ng/ml, and the SCF was added at time 0 at 25 ng/ml. Incubation time was 13 days. The treated cells were then exposed to phorbol myristate acetate (PMA), and Ionomycin. The PMA and Ionomycin were added at 13 days for five hours. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
NERDTDN03	pINCY	This normalized dorsal root ganglion tissue library was constructed from 1.05 million independent clones from a dorsal root ganglion tissue library. Starting RNA was made from dorsal root ganglion tissue removed from the cervical spine of a 32-year-old Caucasian male who died from acute pulmonary edema, acute bronchopneumonia, bilateral pleural effusions, pericardial effusion, and malignant lymphoma (natural killer cell type). The patient presented with pyrexia of unknown origin, malaise, fatigue, and gastrointestinal bleeding. Patient history included probable cytomegalovirus infection, liver congestion, and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, respiratory failure, pneumonia of the left lung, natural killer cell lymphoma of the pharynx, Bell's palsy, and tobacco and alcohol abuse. Previous surgeries included colonoscopy, closed colon biopsy, adenotonsillectomy, and nasopharyngeal endoscopy and biopsy. Patient medications included Diflucan (fluconazole), Deltasone (prednisone), hydrocodone, Lortab, Alprazolam, Reazodone, ProMace-Cytabom, Etoposide, Cisplatin, Cytarabine, and dexamethasone.
		The patient received radiation therapy and multiple blood transfusions. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.

Table 6

Library	Vector	Library Description
NEUTFMT01	PBLUESCRIPT	Library was constructed using total RNA isolated from peripheral blood granulocytes collected by density gradient centrifugation through Ficoll-Hypaque. The cells were isolated from buffy coat units obtained from unrelated male and female donors. Cells were cultured in 10 nM fMLP for 30 minutes, lysed in GuSCN, and spun through CsCl to obtain RNA for library construction. Because this library was made from total RNA, it has an unusually high proportion of unique singleton sequences, which may not all come from polyA RNA species.
NPOLNOT01	pINCY	Library was constructed using RNA isolated from nasal polyp tissue removed from a 78-year-old Caucasian male during a nasal polypectomy. Pathology indicated a nasal polyp and striking eosinophilia. Patient history included asthma and nasal polyps.
OVARTUT02	pINCY	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 51-year-old Caucasian female during an exploratory laparotomy, total abdominal hysterectomy, salpingo-oophorectomy, and an incidental appendectomy. Pathology indicated mucinous cystadenoma presenting as a multiloculated neoplasm involving the entire left ovary. The right ovary contained a follicular cyst and a hemorrhagic corpus luteum. The uterus showed proliferative endometrium and a single intramural leiomyoma. The peritoneal biopsy indicated benign glandular inclusions consistent with endosalpingiosis. Family history included atherosclerotic coronary artery disease, benign hypertension, breast cancer, and uterine cancer.
OVARTUT05	pINCY	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 62-year-old Caucasian female during a total abdominal hysterectomy, removal of the fallopian tubes and ovaries, exploratory laparotomy, regional lymph node excision, and dilation and curettage. Pathology indicated a grade 4 endometrioid carcinoma with extensive squamous differentiation, forming a solid mass in the right ovary. The uterine endometrium was inactive, the cervix showed mild chronic cervicitis, and focal endometriosis was observed in the posterior uterine serosa. Curettings indicated weakly proliferative endometrium with excessive stromal breakdown in the uterus, and a prior cervical biopsy indicated mild chronic cervicitis with a prominent nabothian cyst in the cervix. Patient history included longitudinal deficiency of the radioulna, osteoarthritis, thrombophlebitis, and abnormal blood chemistries. Family history included atherosclerotic coronary artery disease, pulmonary embolism, and cerebrovascular disease.
PANCNOT04	PSPORT1	Library was constructed using RNA isolated from the pancreatic tissue of a 5-year-old Caucasian male who died in a motor vehicle accident.

Table 6

Library	Vector	Library Description
PENITUT01	pINCY	Library was constructed using RNA isolated from tumor tissue removed from the penis of a 64-year-old Caucasian male during penile amputation. Pathology indicated a fungating invasive grade 4 squamous cell carcinoma involving the inner wall of the foreskin and extending onto the glans penis. Patient history included benign neoplasm of the large bowel, atherosclerotic coronary artery disease, angina pectoris, gout, and obesity. Family history included malignant pharyngeal neoplasm, chronic lymphocytic leukemia, and chronic liver disease.
PITUNON01	pINCY	This normalized pituitary gland tissue library was constructed from 6.92 million independent clones from a pituitary gland tissue library. Starting RNA was made from pituitary gland tissue removed from a 55-year-old male who died from chronic obstructive pulmonary disease. Neuropathology indicated there were no gross abnormalities, other than mild ventricular enlargement. There was no apparent microscopic abnormality in any of the neocortical areas examined, except for a number of silver positive neurons with apical dendrite staining, particularly in the frontal lobe. The significance of this was undetermined. The only other microscopic abnormality was that there was prominent silver staining with some swollen axons in the CA3 region of the anterior and posterior hippocampus. Microscopic sections of the cerebellum revealed mild Bergmann's gliosis in the Purkinje cell layer. Patient history included schizophrenia. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
PLACFEB01	pINCY	Library was constructed using pooled cDNA from two different donors. cDNA was generated using RNA isolated from placenta tissue removed from a Caucasian fetus (donor A), who died after 16 weeks' gestation from fetal demise and hydrocephalus; and a Caucasian male fetus (donor B), who died after 18 weeks' gestation from fetal demise. Patient history included umbilical cord wrapped around the head (3 times) and the shoulders (1 time) in donor A. Serology was positive for anti-CMV in donor A. Family history included multiple pregnancies and live births, and an abortion in donor A.
PROSNOT11	pINCY	Library was constructed using RNA isolated from the prostate tissue of a 28-year-old Caucasian male, who died from a self-inflicted gunshot wound.

Table 6

Library	Vector	Library Description
PROSTUS23	pINCY	This subtracted prostate tumor library was constructed using 10 million clones from a pooled prostate tumor library that was subjected to 2 rounds of subtractive hybridization with 10 million clones from a pooled prostate tissue library. The starting library for subtraction was constructed by pooling equal numbers of clones from 4 prostate tumor libraries using mRNA isolated from prostate tumor removed from Caucasian males at ages 58 (A), 61 (B), 66 (C), and 68 (D) during prostatectomy with lymph node excision. Pathology indicated adenocarcinoma in all donors. History included elevated PSA, induration and tobacco abuse in donor A; elevated PSA, induration, prostate hyperplasia, renal failure, osteoarthritis, renal artery stenosis, benign HTN, thrombocytopenia, hyperlipidemia, tobacco/alcohol abuse and hepatitis C (carrier) in donor B; elevated PSA, induration, and tobacco abuse in donor C; and elevated PSA, induration, hypercholesterolemia, and kidney calculus in donor D.
		The hybridization probe for subtraction was constructed by pooling equal numbers of cDNA clones from 3 prostate tissue libraries derived from prostate tissue, prostate epithelial cells, and fibroblasts from prostate stroma from 3 different donors. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR 19 (1991):1954 and Bonaldo, et al. Genome Research 6 (1996):791.
PROTDNV06	PCR2-TOPOTA	Library was constructed using pooled cDNA from different donors. cDNA was generated using mRNA isolated from pooled small intestine tissue removed from a Caucasian male fetus (donor A) who died at 23 weeks' gestation from premature birth; from lung tissue removed from a Caucasian male fetus (donor B) who died from fetal demise; from pleura tumor tissue removed from a 55-year-old Caucasian female (donor C) during a complete pneumonectomy; from frontal/parietal brain tumor tissue removed from a 2-year-old Caucasian female (donor D) during excision of cerebral meningeal lesion; from liver tumor tissue removed from a 72-year-old Caucasian male (donor E) during partial hepatectomy; from pooled fetal brain tissue removed from a Caucasian male fetus (donor F) who was stillborn with a hypoplastic left heart at 23 weeks' gestation and from brain tissue removed from a Caucasian male fetus (donor G), who died at 23 weeks' gestation from premature birth; from pooled fetal kidney tissue removed from 59, 20-33-week-old male and female fetuses who died from spontaneous abortion;

Table 6

Library	Vector	Library Description
		from pooled thymus tissue removed from 9, 18-32-year-old male and female donors who died from sudden death; and from pooled fetal liver tissue removed from 32, 18-24-week-old male and female fetuses. For donor A, serologies were negative. Family history included diabetes in the mother. For donor B, Serologies were negative. For donor C, pathology indicated grade 3 sarcoma most consistent with leiomyosarcoma, uterine primary, forming a bosellated mass replacing the right lower lobe and a portion of the middle lobe. Multiple nodules comprising the tumor show near total necrosis. Smooth muscle actin was positive. Estrogen receptor was negative and progesterone receptor was positive. The patient presented with shortness of breath. Patient history included peptic ulcer disease, normal delivery, anemia, and tobacco abuse in remission. Previous surgeries included total abdominal hysterectomy, bilateral salpingo-oophorectomy, hemorrhoidectomy, endoscopic excision of lung lesion, and appendectomy.
		Patient medications included Megace, tamoxifen, and Pepcid. Family history included multiple sclerosis in the mother; atherosclerotic coronary artery disease and type II diabetes in the father; and breast cancer in the grandparent(s). For donor D, pathology indicated neuroectodermal tumor with advanced ganglionic differentiation. The lesion was only moderately cellular but was mitotically active with a high MIB-1 labelling index. Neuronal differentiation was widespread and advanced. Multinucleate and dysplastic-appearing forms were readily seen. The glial element was less prominent. The patient presented with motor seizures. Family history included hypertension in the grandparent(s). For donor E, pathology indicated metastatic grade 2 (of 4) neuroendocrine carcinoma forming a mass. The patient presented with metastatic liver cancer. Patient history included benign hypertension, type I diabetes, prostatic hyperplasia, prostate cancer, alcohol abuse in remission, and tobacco abuse in remission. Previous surgeries included destruction of a pancreatic lesion, closed prostatic biopsy, transurethral prostatectomy, removal of bilateral testes and total splenectomy.
		Patient medications included Eulexin, Hytrin, Proscar, Ecotrin, and insulin. Family history included atherosclerotic coronary artery disease and acute myocardial infarction in the mother; atherosclerotic coronary artery disease and type II diabetes in the father. For donor F and G, Serologies were negative for both donors and family history for donor G included diabetes in the mother.
SCOMDIC01	PSPORT1	This large size-fractionated library was constructed using RNA isolated from diseased spinal cord tissue removed from the base of the medulla of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Serologies were negative. Patient history included Huntington's disease, emphysema, and tobacco abuse (3-4 packs per day, for 40 years).
SINTNOR01	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 31-year-old Caucasian female during Roux-en-Y gastric bypass. Patient history included clinical obesity.

Table 6

Library	Vector	Library Description
SKINDIA01	PSPORT1	This amplified library was constructed using RNA isolated from diseased skin tissue removed from 1 female and 4 males during skin biopsies. Pathologies indicated tuberculoid and lepromatious leprosy.
SYNORAB01	PBLUESCRIPT	Library was constructed using RNA isolated from the synovial membrane tissue of a 68-year-old Caucasian female with rheumatoid arthritis.
THYRNOT03	pINCY	Library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid. Pathology for the associated tumor tissue indicated dominant follicular adenoma, forming a well-encapsulated mass in the left thyroid.
TLYMNOT08	pINCY	The library was constructed using RNA isolated from anergicallogenic T-lymphocyte tissue removed from an adult (40-50-year old) Caucasian male.The cells were incubated for 3 days in the presence of 1 microgram/ml OKT3 mAb and 5% human serum.
TONSDIT01	pINCY	Library was constructed using RNA isolated from the tonsil tissue of a 6-year-old Caucasian male during adenotonsillectomy. Pathology indicated lymphoid hyperplasia of the tonsils. The patient presented with an abscess of the pharynx. The patient was not taking any medications. Family history included hypothyroidism in the grandparent(s) and benign skin neoplasm in the sibling(s).
UCMCL5T01	PBLUESCRIPT	Library was constructed using RNA isolated from mononuclear cells obtained from the umbilical cord blood of 12 individuals. The cells were cultured for 12 days with IL-5 before RNA was obtained from the pooled lysates.

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value = 1.0E-8 or less; Full Length sequences: Probability value = 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value = 1.06E-6; Assembled ESTs: fasta Identity = 95% or greater and Match length = 200 bases or greater; fastx E value = 1.0E-8 or less; Full Length sequences: fastx score = 100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value = 1.0E-3 or less

Table 7

Program	Description	Reference	Parameter Threshold
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM, INCY, SMART and TIGRFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM, INCY, SMART or TIGRFAM hits: Probability value = 1.0E-3 or less; Signal peptide hits: Score = 0 or greater
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score ≥ GCG specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score = 120 or greater; Match length = 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score = 3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	

Table 7

Program	Description	Reference	Parameter Threshold
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. On Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence (AAAI) Press, Menlo Park, CA, and MIT Press, Cambridge, MA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
59	7313196	011234H1	SNP00074657	59	555	A	A	C	H171	n/d	n/d	n/d	n/d
59	7313196	1287780H1	SNP00065249	110	1177	C	C	T	noncoding	n/a	n/a	n/a	n/a
59	7313196	1315171H1	SNP00128240	270	969	C	C	T	T309	n/a	n/a	n/a	n/a
59	7313196	1739322H1	SNP00074658	179	1007	T	T	G	noncoding	n/d	n/d	n/d	n/d
59	7313196	3095903H1	SNP00013058	81	292	C	C	T	Y83	n/a	n/a	n/a	n/a
59	7313196	6919967H1	SNP00111838	99	636	T	T	G	I198	n/d	n/d	n/d	n/d
59	7313196	6925657H1	SNP00111838	399	630	T	T	G	I196	n/d	n/d	n/d	n/d
60	6465289	1301145H1	SNP00004647	234	4110	T	T	C	I1307	0.89	0.99	0.87	0.94
60	6465289	1301145H1	SNP00108634	188	4063	T	T	C	D1291	n/a	n/a	n/a	n/a
60	6465289	2874757H1	SNP00004646	94	4000	G	G	A	Q1270	n/a	n/a	n/a	n/a
61	7506357	1922848H1	SNP00144854	226	3219	T	T	C	noncoding	n/a	n/a	n/a	n/a
61	7506357	3559382H1	SNP00135225	169	1529	A	A	G	G455	n/a	n/a	n/a	n/a
64	7503356	3405961H1	SNP00101911	37	52	C	T	C	noncoding	n/a	n/a	n/a	n/a
64	7503356	7600509H1	SNP00101911	30	54	T	T	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	001862H1	SNP00058740	268	576	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	002922H1	SNP00076757	102	555	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	003415H1	SNP00058740	45	567	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	010275H1	SNP00067327	5	44	A	A	C	N3	n/a	n/a	n/a	n/a
65	7509052	024132H1	SNP00136036	18	36	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	024145H1	SNP00066370	4	22	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	034545H1	SNP00073392	155	326	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	056145H1	SNP00062326	67	729	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	066164H1	SNP00058740	91	577	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	067517H1	SNP00073454	120	277	G	G	A	noncoding	n/a	n/a	n/a	n/a
65	7509052	075583H1	SNP00073392	151	323	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	083039H1	SNP00066618	9	29	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	100509H1	SNP00067327	27	41	A	A	C	D2	n/a	n/a	n/a	n/a
65	7509052	108973H1	SNP00067327	31	26	A	A	C	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
65	7509052	109780H1	SNP00067327	42	38	A	A	C	K1	n/a	n/a	n/a	n/a
65	7509052	109989H1	SNP00076757	54	556	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	1261380H1	SNP00067327	36	40	A	A	C	I2	n/a	n/a	n/a	n/a
65	7509052	1275449H1	SNP00067327	29	35	A	A	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	133867H1	SNP00073454	92	95	G	G	A	G20	n/a	n/a	n/a	n/a
65	7509052	1369082H1	SNP00076757	40	548	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	1467494H1	SNP00066036	114	280	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	1468472H1	SNP00054956	146	310	G	G	C	noncoding	n/d	n/a	n/a	n/a
65	7509052	147127H1	SNP00067327	28	42	A	A	C	V2	n/a	n/a	n/a	n/a
65	7509052	154011H1	SNP00136036	23	35	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	1543529H1	SNP00066618	18	26	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	155390H1	SNP00136036	17	34	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	1562766H1	SNP00076757	128	552	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	1583035H1	SNP00067327	43	51	A	A	C	T5	n/a	n/a	n/a	n/a
65	7509052	1601118H1	SNP00066036	120	277	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	1612395H1	SNP00136036	29	38	C	C	T	T1	n/a	n/a	n/a	n/a
65	7509052	1624483H1	SNP00066370	12	19	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	1624483H1	SNP00136036	26	33	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	166971H1	SNP00066370	5	18	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	166971H1	SNP00136036	19	31	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	1686007H1	SNP00054956	152	311	G	G	C	noncoding	n/d	n/a	n/a	n/a
65	7509052	1760443H1	SNP00058741	59	721	T	C	T	noncoding	n/d	n/a	n/a	n/a
65	7509052	1911035H1	SNP00058740	17	573	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	1930594H1	SNP00058741	108	724	T	C	T	noncoding	n/d	n/a	n/a	n/a
65	7509052	198049H1	SNP00073454	113	276	G	G	A	noncoding	n/a	n/a	n/a	n/a
65	7509052	2174215H1	SNP00067327	36	43	A	A	C	N3	n/a	n/a	n/a	n/a
65	7509052	2239591H1	SNP00066036	116	276	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	2239916H1	SNP00046811	229	511	C	C	T	noncoding	0.93	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
65	7509052	2239916H1	SNP00092478	47	329	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	2239916H1	SNP00092647	128	410	C	T	C	noncoding	0.15	0.14	0.13	0.10
65	7509052	2287120H1	SNP00092478	38	315	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	2287120H1	SNP00092647	119	396	C	T	C	noncoding	0.15	0.14	0.13	0.10
65	7509052	2325778H1	SNP00076757	186	551	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	2326737H1	SNP00066618	22	30	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	2422724H1	SNP00058741	110	723	C	C	T	noncoding	n/d	n/a	n/a	n/a
65	7509052	2466767H1	SNP00043865	190	390	A	A	G	noncoding	n/d	n/a	n/a	n/a
65	7509052	2493810H1	SNP00066036	108	121	C	C	T	P29	n/a	n/a	n/a	n/a
65	7509052	2550434H1	SNP00066370	13	20	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	2599613H1	SNP00066370	7	17	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	2601252H1	SNP00136036	10	30	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	2640938H1	SNP00066370	10	21	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	2680763H1	SNP00066036	71	120	C	C	T	D28	n/a	n/a	n/a	n/a
65	7509052	274470H1	SNP00066618	16	22	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	2805023H1	SNP00067327	25	39	A	A	C	I1	n/a	n/a	n/a	n/a
65	7509052	2807492H1	SNP00073454	119	275	G	G	A	noncoding	n/a	n/a	n/a	n/a
65	7509052	2815890H1	SNP00067327	22	31	A	A	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	2823643H1	SNP00058740	160	575	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	2824004H1	SNP00066618	7	28	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	2884755H1	SNP00145565	231	485	T	T	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	2895961H1	SNP00066618	11	25	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	2906689H1	SNP00073454	100	107	G	G	A	G24	n/a	n/a	n/a	n/a
65	7509052	2908124H1	SNP00066618	20	24	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	2908738H1	SNP00067327	24	37	A	A	C	M1	n/a	n/a	n/a	n/a
65	7509052	2953302H1	SNP00058740	68	578	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	2959057H1	SNP00058741	50	726	T	C	T	noncoding	n/d	n/a	n/a	n/a
65	7509052	2962649H1	SNP00066618	5	27	T	T	C	noncoding	n/d	n/d	n/d	n/d

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
65	7509052	2964980H1	SNP00067327	50	50	A	A	C	N5	n/a	n/a	n/a	n/a
65	7509052	2989650H1	SNP00067327	32	36	A	A	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	2993573H1	SNP00062326	155	727	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3035151H1	SNP00076757	232	553	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	3081567H1	SNP00073454	83	92	G	G	A	R19	n/a	n/a	n/a	n/a
65	7509052	3106060H1	SNP00076757	166	554	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	3134217H1	SNP00066370	4	16	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3136023H1	SNP00066370	5	12	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3136023H1	SNP00136036	19	26	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3136218H1	SNP00136036	12	32	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	313743H1	SNP00054956	68	307	G	G	C	noncoding	n/d	n/a	n/a	n/a
65	7509052	3139064H1	SNP00005036	13	27	G	G	A	noncoding	n/a	n/a	n/a	n/a
65	7509052	3139117H1	SNP00066036	39	278	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3142293H1	SNP00067327	40	53	A	A	C	E6	n/a	n/a	n/a	n/a
65	7509052	3165712H1	SNP00073454	24	274	G	G	A	noncoding	n/a	n/a	n/a	n/a
65	7509052	3167530H1	SNP00076757	211	549	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	3200709H1	SNP00066370	4	13	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3200709H1	SNP00136036	18	27	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3203056H1	SNP00066370	4	14	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3203056H1	SNP00136036	18	28	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3225449H1	SNP00046811	171	508	T	C	T	noncoding	0.93	n/a	n/a	n/a
65	7509052	3225449H1	SNP00092647	62	399	C	T	C	noncoding	0.15	0.14	0.13	0.10
65	7509052	3233585H1	SNP00067327	27	34	A	A	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	3285021H1	SNP00066618	11	19	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	336496H1	SNP00054956	24	308	G	G	C	noncoding	n/d	n/a	n/a	n/a
65	7509052	3473083H1	SNP00076757	225	550	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	3473389H1	SNP00066618	10	23	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	3481090H1	SNP00054956	250	305	G	G	C	noncoding	n/d	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
65	7509052	3520885H1	SNP00136036	15	24	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3659225H2	SNP00054956	148	309	G	G	C	noncoding	n/d	n/a	n/a	n/a
65	7509052	3661088H1	SNP00092478	4	316	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	3661088H1	SNP00092647	85	397	C	T	C	noncoding	0.15	0.14	0.13	0.10
65	7509052	3808206H1	SNP00066036	100	279	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	385967H1	SNP00046811	218	509	T	C	T	noncoding	0.93	n/a	n/a	n/a
65	7509052	3974169H1	SNP00067327	33	32	A	A	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	4008558H1	SNP00066370	6	15	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	4015892H1	SNP00066036	112	112	C	C	T	R26	n/a	n/a	n/a	n/a
65	7509052	4055824H1	SNP00054956	36	306	G	G	C	noncoding	n/d	n/a	n/a	n/a
65	7509052	4083450H1	SNP00092647	202	402	C	T	C	noncoding	0.15	0.14	0.13	0.10
65	7509052	4085944H1	SNP00136036	6	14	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	4091037H1	SNP00067327	15	33	A	A	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	4145401H1	SNP00073392	68	322	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	4206413H1	SNP00058740	13	574	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	4259219H1	SNP00046811	96	510	T	C	T	noncoding	0.93	n/a	n/a	n/a
65	7509052	4284542H1	SNP00066618	17	20	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	4300710H1	SNP00066618	11	13	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	4310665H1	SNP00066036	101	74	C	C	T	A13	n/a	n/a	n/a	n/a
65	7509052	4312858H1	SNP00103621	97	305	A	A	G	noncoding	0.36	n/a	n/a	n/a
65	7509052	4366064H1	SNP00092647	58	433	C	T	C	noncoding	0.15	0.14	0.13	0.10
65	7509052	4369131H1	SNP00058741	50	727	T	C	T	noncoding	n/d	n/a	n/a	n/a
65	7509052	4381549H2	SNP00062326	107	728	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	446350H1	SNP00062326	162	721	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	4664150H1	SNP00136036	13	22	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	4694537H1	SNP00145565	143	483	T	T	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	4726131H1	SNP00092647	59	398	C	T	C	noncoding	0.15	0.14	0.13	0.10
65	7509052	4829345H1	SNP00066618	13	17	T	T	C	noncoding	n/d	n/d	n/d	n/d

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
65	7509052	4832194H1	SNP00136036	20	23	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	4975571H1	SNP00136036	18	29	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	5084961H1	SNP00149316	160	537	G	G	A	noncoding	n/a	n/a	n/a	n/a
65	7509052	5111296H1	SNP00092478	109	373	T	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	5170779H1	SNP00066036	105	106	C	C	T	L24	n/a	n/a	n/a	n/a
65	7509052	5200251H1	SNP00136036	9	13	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	5211977H1	SNP00073454	97	104	G	G	A	G23	n/a	n/a	n/a	n/a
65	7509052	5222143H2	SNP00073454	24	38	G	G	A	R1	n/a	n/a	n/a	n/a
65	7509052	5282009H1	SNP00062326	199	724	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	5282723H1	SNP00073392	242	327	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	5438072H1	SNP00067327	29	47	A	A	C	H4	n/a	n/a	n/a	n/a
65	7509052	5559509H1	SNP00073454	120	118	G	G	A	D28	n/a	n/a	n/a	n/a
65	7509052	5576233H1	SNP00066036	84	93	C	C	T	R19	n/a	n/a	n/a	n/a
65	7509052	5576233H1	SNP00067327	5	15	A	A	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	5608711H1	SNP00058740	192	579	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	5668040H1	SNP00066036	127	285	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	5669353H1	SNP00067327	42	45	A	A	C	K3	n/a	n/a	n/a	n/a
65	7509052	5671107H1	SNP00066036	102	105	C	C	T	D23	n/a	n/a	n/a	n/a
65	7509052	5671107H1	SNP00067327	23	25	A	A	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	5765576H1	SNP00073454	94	117	G	G	A	L27	n/a	n/a	n/a	n/a
65	7509052	578350H1	SNP00066618	15	18	T	T	C	noncoding	n/d	n/d	n/d	n/d
65	7509052	580595H1	SNP00067327	30	29	A	A	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	5878426H1	SNP00136036	8	19	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	5911743H1	SNP00076757	54	537	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	5911858H1	SNP00073454	97	116	G	G	A	R27	n/a	n/a	n/a	n/a
65	7509052	5941194H1	SNP00066036	20	21	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	604424H1	SNP00076757	257	545	G	G	A	noncoding	n/d	n/d	n/d	n/d
65	7509052	6075860H1	SNP00066618	11	21	T	T	C	noncoding	n/d	n/d	n/d	n/d

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
65	7509052	6092530H1	SNP00066370	2	11	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	6092530H1	SNP00136036	16	25	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	6104923H1	SNP00136036	12	21	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	615004H1	SNP00066036	131	290	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	6154318H1	SNP00066036	105	107	C	C	T	A24	n/a	n/a	n/a	n/a
65	7509052	6154318H1	SNP00067327	26	28	A	A	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	6399289H1	SNP00066036	105	271	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	640291H1	SNP00066370	19	23	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	640291H1	SNP00136036	33	37	C	C	T	L1	n/a	n/a	n/a	n/a
65	7509052	6412005H1	SNP00110910	396	274	T	T	G	noncoding	n/d	n/a	n/a	n/a
65	7509052	6414114H1	SNP00073392	161	320	G	G	C	noncoding	n/a	n/a	n/a	n/a
65	7509052	6425507H1	SNP00054956	33	38	G	G	C	R1	n/d	n/a	n/a	n/a
65	7509052	646225H1	SNP00066370	27	26	C	C	T	noncoding	n/a	n/a	n/a	n/a
65	7509052	646225H1	SNP00136036	41	40	C	C	T	L2	n/a	n/a	n/a	n/a
65	7509052	6531208H1	SNP00058741	208	734	T	C	T	noncoding	n/d	n/a	n/a	n/a
66	7503366	1379965H1	SNP00114801	155	3793	A	A	G	noncoding	n/a	n/a	n/a	n/a
66	7503366	2632730H1	SNP00146109	157	2206	A	A	G	noncoding	n/a	n/a	n/a	n/a
66	7503366	2758103H1	SNP00019895	76	3067	T	C	T	noncoding	n/d	n/d	n/d	n/d
66	7503366	3090177H1	SNP00019895	250	3066	C	C	T	noncoding	n/d	n/d	n/d	n/d
66	7503366	3438044H1	SNP00114801	156	3791	A	A	G	noncoding	n/a	n/a	n/a	n/a
66	7503366	5642352H1	SNP00019895	250	3069	C	C	T	noncoding	n/d	n/d	n/d	n/d
67	7505933	1605914H1	SNP00067454	26	630	C	C	G	A194	n/d	n/a	n/a	n/a
67	7505933	1627903H1	SNP00112571	72	1201	A	A	C	noncoding	n/d	n/a	n/a	n/a
67	7505933	1923831H1	SNP00067455	160	994	C	C	T	V315	n/a	n/a	n/a	n/a
67	7505933	1959225H1	SNP00067455	62	989	C	C	T	R314	n/a	n/a	n/a	n/a
67	7505933	1959225H1	SNP00067456	181	1108	G	G	C	K353	n/a	n/a	n/a	n/a
67	7505933	2120640H1	SNP00147603	101	1008	C	C	G	P320	n/a	n/a	n/a	n/a
67	7505933	2330059H1	SNP00067456	15	1113	G	G	C	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
67	7505933	3704244H1	SNP00067454	62	629	C	C	G	P194	n/d	n/a	n/a	n/a
67	7505933	3838492H1	SNP00067455	246	992	C	C	T	L315	n/a	n/a	n/a	n/a
67	7505933	4635765H1	SNP00067456	120	1111	G	G	C	K354	n/a	n/a	n/a	n/a
67	7505933	4698117H1	SNP00067455	129	993	C	C	T	A315	n/a	n/a	n/a	n/a
67	7505933	4698117H1	SNP00067456	248	1112	G	G	C	noncoding	n/a	n/a	n/a	n/a
67	7505933	4802365H1	SNP00147603	31	1005	C	C	G	P319	n/a	n/a	n/a	n/a
67	7505933	5949929H1	SNP00113220	250	726	G	G	A	G226	n/a	n/a	n/a	n/a
67	7505933	6100378H1	SNP00147603	123	1014	C	C	G	A322	n/a	n/a	n/a	n/a
67	7505933	752417H1	SNP00067454	39	628	C	C	G	L193	n/d	n/a	n/a	n/a
67	7505933	832728H1	SNP00067456	31	1115	G	G	C	noncoding	n/a	n/a	n/a	n/a
70	2008979	1695989H1	SNP00106662	102	809	T	T	C	L252	n/a	n/a	n/a	n/a
70	2008979	2622589H1	SNP00106662	167	803	T	T	C	V250	n/a	n/a	n/a	n/a
70	2008979	2954591H1	SNP00106663	239	1601	A	A	G	N516	n/d	n/d	n/d	n/d
70	2008979	3799293H1	SNP00106663	205	1598	A	A	G	stop515	n/d	n/d	n/d	n/d
70	2008979	3800193H1	SNP00106663	206	1600	A	A	G	N516	n/d	n/d	n/d	n/d
70	2008979	4401805H1	SNP00106660	38	508	A	A	G	N152	n/a	n/a	n/a	n/a
70	2008979	4553835H1	SNP00106663	45	1599	A	A	G	L515	n/d	n/d	n/d	n/d
70	2008979	4592620H1	SNP00132319	124	936	A	A	G	K294	n/a	n/a	n/a	n/a
70	2008979	4875544H1	SNP00106661	38	547	A	A	G	T165	n/a	n/a	n/a	n/a
70	2008979	4875544H1	SNP00106662	292	801	T	T	C	L249	n/a	n/a	n/a	n/a
70	2008979	5168928H1	SNP00106661	23	546	A	A	G	L164	n/a	n/a	n/a	n/a
70	2008979	5427451H1	SNP00132319	166	938	A	A	G	D295	n/a	n/a	n/a	n/a
70	2008979	6528070H1	SNP00106660	145	509	A	A	G	Y152	n/a	n/a	n/a	n/a
70	2008979	6528070H1	SNP00106661	185	549	A	A	G	A165	n/a	n/a	n/a	n/a
71	90073157	1352667H1	SNP00010549	41	412	T	T	C	noncoding	0.46	0.45	0.33	0.35
71	90073157	182066H1	SNP00010549	111	411	T	T	C	noncoding	0.46	0.45	0.33	0.35
71	90073157	4539814H1	SNP00099284	85	633	T	T	C	noncoding	n/a	n/a	n/a	n/a
71	90073157	4590421H1	SNP00010549	100	409	C	T	C	noncoding	0.46	0.45	0.33	0.35

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
72	7506782	1486436H1	SNP00036529	120	43	C	C	G	F7	0.91	0.83	0.94	0.85
72	7506782	1486436H1	SNP00036530	150	73	A	A	G	T17	n/a	n/a	n/a	n/a
72	7506782	2400259H1	SNP00036531	64	1257	T	T	C	noncoding	0.99	n/d	n/d	n/d
72	7506782	2534038H1	SNP00036529	114	44	C	C	G	P8	0.91	0.83	0.94	0.85
72	7506782	2534038H1	SNP00036530	144	74	A	A	G	R18	n/a	n/a	n/a	n/a
72	7506782	4309005H1	SNP00036530	123	60	A	A	G	Q13	n/a	n/a	n/a	n/a
72	7506782	4548843H1	SNP00036529	113	39	C	C	G	P6	0.91	0.83	0.94	0.85
72	7506782	4548843H1	SNP00036530	142	69	A	A	G	D16	n/a	n/a	n/a	n/a
72	7506782	4761388H1	SNP00036529	94	26	C	C	G	P2	0.91	0.83	0.94	0.85
72	7506782	5204618H1	SNP00036531	214	1255	T	T	C	noncoding	0.99	n/d	n/d	n/d
72	7506782	6398707H1	SNP00036531	110	1253	T	T	C	noncoding	0.99	n/d	n/d	n/d
72	7506782	6740381H1	SNP00036531	209	1051	A	A	G	E343	0.99	n/d	n/d	n/d
72	7506782	7054763H1	SNP00036531	32	1050	A	A	G	E343	0.99	n/d	n/d	n/d
73	7506941	1689806H1	SNP00019669	37	2163	C	C	T	noncoding	n/a	n/a	n/a	n/a
73	7506941	5444612H1	SNP00019669	151	1382	C	C	T	T256	n/a	n/a	n/a	n/a
73	7506941	7268478H1	SNP00114227	111	1449	C	T	C	P278	n/d	n/a	n/a	n/a
74	7507072	1328401H1	SNP00034034	65	988	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	1328401H1	SNP00046549	23	946	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	1330788H1	SNP00034034	100	989	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	1330788H1	SNP00046549	142	947	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	1333184H1	SNP00034034	60	992	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	1333184H1	SNP00046549	18	950	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	1462605H1	SNP00034034	51	985	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	1462605H1	SNP00046549	9	943	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	1462893H1	SNP00046549	50	949	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	1529202H1	SNP00034034	136	990	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	1529202H1	SNP00046549	94	948	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	2068958H1	SNP00034034	131	993	C	C	T	noncoding	n/d	n/d	n/d	n/d

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
74	7507072	2068958H1	SNP00046549	89	951	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	2087314H1	SNP00034034	111	1006	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	2087314H1	SNP00046549	69	965	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	2090187H1	SNP00034034	145	991	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	2093817H1	SNP00034034	64	987	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	2093817H1	SNP00046549	22	945	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	2725281H1	SNP00034954	229	854	T	T	C	noncoding	0.87	0.84	n/d	0.97
74	7507072	3163982H1	SNP00034954	19	853	C	T	C	noncoding	0.87	0.84	n/d	0.97
74	7507072	4170095H1	SNP00046549	110	999	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	4220237H1	SNP00034034	81	1003	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	4220237H1	SNP00046549	39	961	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	948556H1	SNP00034034	127	1000	C	C	T	noncoding	n/d	n/d	n/d	n/d
74	7507072	948556H1	SNP00046549	85	958	C	C	T	noncoding	n/d	n/d	n/d	n/d
76	7509097	1471480H1	SNP00063744	129	913	T	T	C	noncoding	n/a	n/a	n/a	n/a
77	7509118	1471480H1	SNP00063744	129	1115	T	T	C	S368	n/a	n/a	n/a	n/a
78	7509312	1471480H1	SNP00063744	129	952	T	T	C	S302	n/a	n/a	n/a	n/a
79	90126902	7655009J1	SNP00112109	368	1856	A	A	G	noncoding	n/a	n/a	n/a	n/a
79	90126902	7680089J1	SNP00112109	286	2145	A	A	G	noncoding	n/a	n/a	n/a	n/a
80	7509352	2546514H2	SNP00019330	47	1018	C	T	C	noncoding	0.28	0.18	0.47	0.29
80	7509352	2782888H2	SNP00019330	18	1019	T	T	C	noncoding	0.28	0.18	0.47	0.29
80	7509352	3731340H1	SNP00019330	151	1016	C	T	C	noncoding	0.28	0.18	0.47	0.29
83	7500455	1354192H1	SNP00010952	76	3905	C	C	T	noncoding	n/d	n/a	n/a	n/a
83	7500455	1354192H1	SNP00041608	106	3935	C	C	G	noncoding	n/a	n/a	n/a	n/a
83	7500455	1609808H1	SNP00061317	170	3280	A	A	G	noncoding	n/d	n/d	n/d	n/d
83	7500455	2006213H1	SNP00061318	151	3457	C	C	T	noncoding	n/a	n/a	n/a	n/a
83	7500455	2183868H1	SNP00041607	224	2818	C	C	T	noncoding	n/a	n/a	n/a	n/a
83	7500455	2285953H1	SNP00141178	159	3542	T	T	C	noncoding	n/a	n/a	n/a	n/a
83	7500455	2346489H1	SNP00041605	185	1482	G	G	A	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
83	7500455	2346489H1	SNP00041606	191	1488	G	G	A	noncoding	n/a	n/a	n/a	n/a
83	7500455	2644846H1	SNP00061316	38	1992	C	C	A	noncoding	n/a	n/a	n/a	n/a
83	7500455	2848792H1	SNP00061317	244	3276	A	A	G	noncoding	n/d	n/d	n/d	n/d
83	7500455	2851283H1	SNP00061318	11	3454	C	C	T	noncoding	n/a	n/a	n/a	n/a
83	7500455	2851283H1	SNP00141178	96	3539	T	T	C	noncoding	n/a	n/a	n/a	n/a
83	7500455	2868922H1	SNP00041607	138	2814	C	C	T	noncoding	n/a	n/a	n/a	n/a
83	7500455	3041688H1	SNP00061317	131	3281	A	A	G	noncoding	n/d	n/d	n/d	n/d
83	7500455	3376953H1	SNP00041607	21	2816	C	C	T	noncoding	n/a	n/a	n/a	n/a
83	7500455	3441344H1	SNP00061317	142	3279	A	A	G	noncoding	n/d	n/d	n/d	n/d
83	7500455	3607107H1	SNP00061318	129	3451	C	C	T	noncoding	n/a	n/a	n/a	n/a
83	7500455	3607107H1	SNP00141178	214	3536	T	T	C	noncoding	n/a	n/a	n/a	n/a
83	7500455	3691013H1	SNP00061318	36	3455	C	C	T	noncoding	n/a	n/a	n/a	n/a
83	7500455	3691013H1	SNP00141178	121	3540	T	T	C	noncoding	n/a	n/a	n/a	n/a
83	7500455	4419530H1	SNP00041607	64	2815	C	C	T	noncoding	n/a	n/a	n/a	n/a
83	7500455	4640760H1	SNP00061317	258	3278	A	A	G	noncoding	n/d	n/d	n/d	n/d
83	7500455	4646626H1	SNP00061317	131	3268	A	A	G	noncoding	n/d	n/d	n/d	n/d
83	7500455	4651554H1	SNP00041607	88	2817	C	C	T	noncoding	n/a	n/a	n/a	n/a
83	7500455	5497751H1	SNP00149382	59	1095	C	C	A	noncoding	n/a	n/a	n/a	n/a
83	7500455	5508321H1	SNP00129654	71	98	T	T	C	noncoding	n/a	n/a	n/a	n/a
83	7500455	6079980H1	SNP00061315	214	1089	G	G	A	noncoding	0.95	0.99	n/d	0.97
83	7500455	6438022H1	SNP00135482	280	1901	T	T	C	noncoding	n/a	n/a	n/a	n/a
83	7500455	953811H1	SNP00141178	155	3541	T	T	C	noncoding	n/a	n/a	n/a	n/a
84	7510401	1458452H1	SNP00100985	43	1198	A	A	C	noncoding	n/a	n/a	n/a	n/a
84	7510401	1980663H1	SNP00021508	32	500	A	C	A	S109	n/d	n/d	n/d	n/d
84	7510401	3121411H1	SNP00127206	30	1094	A	A	G	noncoding	n/a	n/a	n/a	n/a
84	7510401	877913H1	SNP00143628	37	1111	C	T	C	noncoding	n/a	n/a	n/a	n/a
85	7504702	1240439H1	SNP00011014	51	646	T	T	C	noncoding	n/a	n/a	n/a	n/a
85	7504702	1240439H1	SNP00041706	214	809	G	G	A	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
85	7504702	1242894H1	SNP00011013	68	124	T	T	C	I40	0.68	n/a	n/a	n/a
85	7504702	1359840H1	SNP00074796	18	451	T	T	G	noncoding	n/a	n/a	n/a	n/a
85	7504702	1361523H1	SNP00011014	100	644	T	T	C	noncoding	n/a	n/a	n/a	n/a
85	7504702	1749008H1	SNP00023897	77	645	C	C	G	noncoding	n/d	n/d	0.97	n/d
85	7504702	1984153H1	SNP00041706	253	808	G	G	A	noncoding	n/a	n/a	n/a	n/a
85	7504702	2435410H1	SNP00153006	111	704	T	G	T	noncoding	n/a	n/a	n/a	n/a
85	7504702	3342691H1	SNP00023897	56	642	C	C	G	noncoding	n/d	n/d	0.97	n/d
85	7504702	3342691H1	SNP00153006	115	701	G	G	T	noncoding	n/a	n/a	n/a	n/a
85	7504702	3387137H1	SNP00011014	209	647	T	T	C	noncoding	n/a	n/a	n/a	n/a
85	7504702	3387984H1	SNP00041705	107	163	C	T	C	P53	n/a	n/a	n/a	n/a
85	7504702	3481762H1	SNP00011014	136	645	T	T	C	noncoding	n/a	n/a	n/a	n/a
85	7504702	3643951H1	SNP00041705	106	161	T	T	C	D52	n/a	n/a	n/a	n/a
85	7504702	3645917H1	SNP00011013	68	89	T	T	C	R28	0.68	n/a	n/a	n/a
85	7504702	3646351H1	SNP00041705	107	162	T	T	C	F53	n/a	n/a	n/a	n/a
85	7504702	3979308H1	SNP00011013	68	125	T	T	C	I40	0.68	n/a	n/a	n/a
85	7504702	4437418H1	SNP00011014	86	641	T	T	C	noncoding	n/a	n/a	n/a	n/a
85	7504702	5022209H1	SNP00041706	232	807	G	G	A	noncoding	n/a	n/a	n/a	n/a
85	7504702	5551168H1	SNP00074796	77	425	G	T	G	noncoding	n/a	n/a	n/a	n/a
85	7504702	877669H1	SNP00074796	59	453	T	T	G	noncoding	n/a	n/a	n/a	n/a
86	7509113	1815032H1	SNP00122239	218	621	C	C	A	T189	n/d	n/a	n/a	n/a
86	7509113	1849951H1	SNP00011974	113	145	G	G	C	Q30	n/d	n/a	n/a	n/a
86	7509113	1987840H1	SNP00011974	122	146	G	G	C	A31	n/d	n/a	n/a	n/a
86	7509113	2651951H1	SNP00044624	133	758	C	C	T	H235	n/a	n/a	n/a	n/a
86	7509113	2820905H1	SNP00011974	85	142	G	G	C	E29	n/d	n/a	n/a	n/a
86	7509113	3036853H1	SNP00044623	151	715	T	T	C	T220	n/a	n/a	n/a	n/a
86	7509113	3036853H1	SNP00044624	193	757	T	C	T	L234	n/a	n/a	n/a	n/a
86	7509113	3201751H1	SNP00011974	135	143	G	G	C	E30	n/d	n/a	n/a	n/a
86	7509113	3230305H1	SNP00011974	63	144	G	G	C	R30	n/d	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
86	7509113	4762858H1	SNP00044623	65	713	T	T	C	S220	n/a	n/a	n/a	n/a
86	7509113	4762858H1	SNP00044624	107	755	C	C	T	L234	n/a	n/a	n/a	n/a
86	7509113	6163261H1	SNP00052452	515	1121	C	C	T	Q356	n/d	n/d	n/d	n/d
87	7509140	1815032H1	SNP00122239	218	743	C	C	A	T233	n/d	n/a	n/a	n/a
87	7509140	1849951H1	SNP00011974	113	135	G	G	C	Q30	n/d	n/a	n/a	n/a
87	7509140	1987840H1	SNP00011974	122	136	G	G	C	A31	n/d	n/a	n/a	n/a
87	7509140	2820905H1	SNP00011974	85	132	G	G	C	E29	n/d	n/a	n/a	n/a
87	7509140	3201751H1	SNP00011974	135	133	G	G	C	E30	n/d	n/a	n/a	n/a
87	7509140	3230305H1	SNP00011974	63	134	G	G	C	R30	n/d	n/a	n/a	n/a
87	7509140	6207548H1	SNP00052452	314	910	T	C	T	stop289	n/d	n/d	n/d	n/d
87	7509140	6308751H1	SNP00011974	109	129	G	G	C	R28	n/d	n/a	n/a	n/a
88	7509223	1294161H1	SNP00132638	219	478	C	C	T	noncoding	n/a	n/a	n/a	n/a
88	7509223	1294161H1	SNP00132639	251	446	G	G	A	noncoding	n/a	n/a	n/a	n/a
88	7509223	1509416H1	SNP00132640	161	193	T	T	C	noncoding	n/a	n/a	n/a	n/a
88	7509223	179220H1	SNP00132639	88	444	G	G	A	noncoding	n/a	n/a	n/a	n/a
88	7509223	1815032H1	SNP00122239	218	1324	C	C	A	T233	n/d	n/a	n/a	n/a
88	7509223	1849951H1	SNP00011974	113	716	G	G	C	Q30	n/d	n/a	n/a	n/a
88	7509223	1987840H1	SNP00011974	122	717	G	G	C	A31	n/d	n/a	n/a	n/a
88	7509223	2696377H1	SNP00132638	28	483	C	C	T	noncoding	n/a	n/a	n/a	n/a
88	7509223	2696377H1	SNP00132639	60	451	G	G	A	noncoding	n/a	n/a	n/a	n/a
88	7509223	2820905H1	SNP00011974	85	713	G	G	C	E29	n/d	n/a	n/a	n/a
88	7509223	2864068H1	SNP00132638	226	480	C	C	T	noncoding	n/a	n/a	n/a	n/a
88	7509223	3140352H1	SNP00151210	188	420	A	A	G	noncoding	n/a	n/a	n/a	n/a
88	7509223	3201751H1	SNP00011974	135	714	G	G	C	E30	n/d	n/a	n/a	n/a
88	7509223	3206181H1	SNP00132638	34	481	C	C	T	noncoding	n/a	n/a	n/a	n/a
88	7509223	3206181H1	SNP00132639	66	449	G	G	A	noncoding	n/a	n/a	n/a	n/a
88	7509223	3230305H1	SNP00011974	63	715	G	G	C	R30	n/d	n/a	n/a	n/a
88	7509223	3542643H1	SNP00132639	11	448	G	G	A	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
88	7509223	3702420H1	SNP00132638	80	479	C	C	T	noncoding	n/a	n/a	n/a	n/a
88	7509223	3702420H1	SNP00132639	112	447	G	G	A	noncoding	n/a	n/a	n/a	n/a
88	7509223	3705854H1	SNP00132640	219	194	T	T	C	noncoding	n/a	n/a	n/a	n/a
88	7509223	3811445H1	SNP00132640	129	197	T	T	C	noncoding	n/a	n/a	n/a	n/a
88	7509223	3879778H1	SNP00132640	119	202	T	T	C	noncoding	n/a	n/a	n/a	n/a
88	7509223	4357880H1	SNP00151210	53	426	A	A	G	noncoding	n/a	n/a	n/a	n/a
88	7509223	4561296H1	SNP00132640	180	204	T	T	C	noncoding	n/a	n/a	n/a	n/a
88	7509223	4640080H1	SNP00132638	107	482	C	C	T	noncoding	n/a	n/a	n/a	n/a
88	7509223	4640080H1	SNP00132639	139	450	G	G	A	noncoding	n/a	n/a	n/a	n/a
88	7509223	5868577H1	SNP00132640	183	229	T	T	C	noncoding	n/a	n/a	n/a	n/a
88	7509223	5954870H1	SNP00132640	128	196	T	T	C	noncoding	n/a	n/a	n/a	n/a
88	7509223	6030484H1	SNP00132640	98	195	T	T	C	noncoding	n/a	n/a	n/a	n/a
88	7509223	6146025H1	SNP00151210	367	419	A	A	G	noncoding	n/a	n/a	n/a	n/a
88	7509223	6183919H1	SNP00131425	12	108	A	A	G	noncoding	n/a	n/a	n/a	n/a
88	7509223	6207548H1	SNP00052452	314	1686	T	C	T	stop354	n/d	n/d	n/d	n/d
88	7509223	6260421H1	SNP00131425	233	109	A	A	G	noncoding	n/a	n/a	n/a	n/a
88	7509223	6308751H1	SNP00011974	109	710	G	G	C	R28	n/d	n/a	n/a	n/a
88	7509223	7174182H2	SNP00132640	273	172	T	T	C	noncoding	n/a	n/a	n/a	n/a
88	7509223	8616728J1	SNP00132639	73	426	G	G	A	noncoding	n/a	n/a	n/a	n/a
88	7509223	8616728J1	SNP00132640	327	174	T	T	C	noncoding	n/a	n/a	n/a	n/a
89	7509272	1471480H1	SNP00063744	129	974	T	T	C	S302	n/a	n/a	n/a	n/a
90	7509327	1471480H1	SNP00063744	129	841	T	T	C	noncoding	n/a	n/a	n/a	n/a
92	7504534	1538120H1	SNP00024718	40	2633	C	C	T	S871	n/a	n/a	n/a	n/a
92	7504534	2014759H1	SNP00074271	137	2523	A	A	G	R835	n/d	n/d	n/d	n/d
92	7504534	2602020H1	SNP00122174	256	1590	G	G	C	G524	n/a	n/a	n/a	n/a
92	7504534	3248601H1	SNP00122174	104	1592	G	G	C	R524	n/a	n/a	n/a	n/a
92	7504534	3777007H1	SNP00123535	160	182	A	A	G	stop54	0.99	n/d	n/d	n/d
92	7504534	3796741H1	SNP00024718	250	2632	C	C	T	S871	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
92	7504534	3796741H1	SNP00074271	140	2522	A	A	G	stop834	n/d	n/d	n/d	n/d
92	7504534	4548418H1	SNP00024718	247	2631	C	C	T	P871	n/a	n/a	n/a	n/a
92	7504534	4548418H1	SNP00074271	137	2521	A	A	G	stop834	n/d	n/d	n/d	n/d
92	7504534	4814965H1	SNP00149887	84	2575	C	C	T	S852	n/a	n/a	n/a	n/a
92	7504534	5044629H1	SNP00136617	83	458	T	T	C	S146	n/a	n/a	n/a	n/a
92	7504534	6095442H1	SNP00122174	275	1589	G	G	C	S523	n/a	n/a	n/a	n/a
92	7504534	6144495H1	SNP00149887	132	2576	C	C	T	S852	n/a	n/a	n/a	n/a
92	7504534	6202336H1	SNP00123535	161	184	A	A	G	D55	0.99	n/d	n/d	n/d
92	7504534	6448564H1	SNP00074271	287	2515	A	A	G	K832	n/d	n/d	n/d	n/d
93	7507771	2116838H1	SNP00129102	125	10730	A	A	G	Y3424	n/a	n/a	n/a	n/a
93	7507771	2121221H1	SNP00017169	11	11533	A	A	G	I3692	n/d	n/a	n/a	n/a
93	7507771	3298296H1	SNP00017169	190	11531	A	A	G	E3691	n/d	n/a	n/a	n/a
93	7507771	3563929H1	SNP00053087	129	9653	T	C	T	L3065	n/d	n/a	n/a	n/a
93	7507771	3603076H1	SNP00017169	51	11530	A	A	G	I3691	n/d	n/a	n/a	n/a
93	7507771	6046434H1	SNP00053087	412	9656	C	C	T	P3066	n/d	n/a	n/a	n/a
93	7507771	6144483H1	SNP00112143	236	1907	A	A	G	N483	0.97	n/d	n/d	n/d
93	7507771	6420651H1	SNP00129102	240	10669	A	A	G	S3404	n/a	n/a	n/a	n/a
93	7507771	6804640J1	SNP00109051	178	7522	C	T	C	L2355	n/a	n/a	n/a	n/a
94	7504732	654991H1	SNP00141070	78	1224	C	C	T	noncoding	n/a	n/a	n/a	n/a
96	7459720	1270817H1	SNP00007805	12	2551	A	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	1270817H1	SNP00107784	72	2610	G	G	A	noncoding	n/d	n/a	n/a	n/a
96	7459720	1345356H1	SNP00033582	87	102	G	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	1562910H1	SNP00062531	69	2518	C	C	T	P783	n/a	n/a	n/a	n/a
96	7459720	1675153H1	SNP00004174	132	2225	T	T	C	T685	n/d	n/d	0.21	0.43
96	7459720	1731831H1	SNP00108561	131	860	G	A	G	T230	0.16	0.44	0.12	0.34
96	7459720	1867507H1	SNP00007805	75	2547	A	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	1867507H1	SNP00062531	108	2514	C	C	T	L782	n/a	n/a	n/a	n/a
96	7459720	1867507H1	SNP00107784	15	2606	G	G	A	noncoding	n/d	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
96	7459720	2099327H1	SNP00057034	120	2099	C	C	T	Y643	n/a	n/a	n/a	n/a
96	7459720	2099327H1	SNP00108563	197	2176	A	A	G	N669	n/a	n/a	n/a	n/a
96	7459720	2262828H1	SNP00004174	85	2224	T	T	C	I685	n/d	n/d	0.21	0.43
96	7459720	2276094H1	SNP00007805	67	2550	G	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	2276094H1	SNP00062531	100	2517	C	C	T	P783	n/a	n/a	n/a	n/a
96	7459720	2276094H1	SNP00107784	7	2609	G	G	A	noncoding	n/d	n/a	n/a	n/a
96	7459720	2436233H1	SNP00076056	168	1038	C	C	T	L290	n/a	n/a	n/a	n/a
96	7459720	2448954H1	SNP00033582	83	103	A	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	2654420H1	SNP00057034	146	2088	C	C	T	R640	n/a	n/a	n/a	n/a
96	7459720	2654420H1	SNP00108563	223	2165	A	A	G	T665	n/a	n/a	n/a	n/a
96	7459720	3088462H1	SNP00007805	150	2506	A	A	G	H779	n/a	n/a	n/a	n/a
96	7459720	3088462H1	SNP00062531	183	2472	C	C	T	Q768	n/a	n/a	n/a	n/a
96	7459720	3088462H1	SNP00107784	90	2567	G	G	A	noncoding	n/d	n/a	n/a	n/a
96	7459720	3168491H1	SNP00004174	148	2223	C	T	C	P685	n/d	n/d	0.21	0.43
96	7459720	3207088H1	SNP00107784	140	2573	G	G	A	noncoding	n/d	n/a	n/a	n/a
96	7459720	3513937H1	SNP00007805	70	2531	A	A	G	E787	n/a	n/a	n/a	n/a
96	7459720	3513937H1	SNP00062531	103	2498	C	C	T	P776	n/a	n/a	n/a	n/a
96	7459720	3674609H1	SNP00057033	235	1953	T	T	G	W595	0.90	n/a	n/a	n/a
96	7459720	3961677H1	SNP00007805	96	2484	A	A	G	T772	n/a	n/a	n/a	n/a
96	7459720	3961677H1	SNP00062531	129	2451	C	C	T	Q761	n/a	n/a	n/a	n/a
96	7459720	4114958H1	SNP00007805	159	2495	G	A	G	W775	n/a	n/a	n/a	n/a
96	7459720	4114958H1	SNP00062531	192	2462	T	C	T	C764	n/a	n/a	n/a	n/a
96	7459720	4440675H1	SNP00062531	122	2516	C	C	T	V782	n/a	n/a	n/a	n/a
96	7459720	4500882H1	SNP00033582	91	94	A	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	4600141H1	SNP00007805	13	2546	A	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	4600141H1	SNP00107784	73	2605	G	G	A	noncoding	n/d	n/a	n/a	n/a
96	7459720	4797817H1	SNP00007805	100	2549	A	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	4797817H1	SNP00107784	160	2608	G	G	A	noncoding	n/d	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
96	7459720	4851201H1	SNP00007805	202	2552	A	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	4851201H1	SNP00107784	142	2611	G	G	A	noncoding	n/d	n/a	n/a	n/a
96	7459720	5100026H1	SNP00007805	138	2544	A	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	5100026H1	SNP00062531	171	2511	C	C	T	H781	n/a	n/a	n/a	n/a
96	7459720	5100026H1	SNP00107784	78	2603	G	G	A	noncoding	n/d	n/a	n/a	n/a
96	7459720	5662941H1	SNP00057034	43	2098	C	C	T	S643	n/a	n/a	n/a	n/a
96	7459720	5662941H1	SNP00108563	120	2175	A	A	G	N669	n/a	n/a	n/a	n/a
96	7459720	5677175H1	SNP00033582	62	75	A	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	6422535H1	SNP00033582	71	84	A	A	G	noncoding	n/a	n/a	n/a	n/a
96	7459720	6610009H1	SNP00108562	115	1023	C	T	C	R285	n/a	n/a	n/a	n/a
96	7459720	6798811J1	SNP00076056	46	1022	C	C	T	Y284	n/a	n/a	n/a	n/a
96	7459720	6814531H1	SNP00057033	356	1958	T	T	G	L596	0.90	n/a	n/a	n/a
96	7459720	6938337H1	SNP00007805	171	2538	A	A	G	I790	n/a	n/a	n/a	n/a
96	7459720	6938337H1	SNP00062531	138	2505	C	C	T	H779	n/a	n/a	n/a	n/a
96	7459720	6938337H1	SNP00107784	231	2597	G	G	A	noncoding	n/d	n/a	n/a	n/a
96	7459720	6949559H1	SNP00108562	381	1041	C	T	C	L291	n/a	n/a	n/a	n/a
96	7459720	7402138H1	SNP00108562	43	1025	T	T	C	S285	n/a	n/a	n/a	n/a
97	7503300	1348487H1	SNP00114445	182	1171	G	G	A	R340	n/a	n/a	n/a	n/a
97	7503300	1495867H1	SNP00044495	39	1843	T	C	T	noncoding	n/a	n/a	n/a	n/a
97	7503300	4595762H1	SNP00114445	252	1168	G	G	A	R339	n/a	n/a	n/a	n/a
98	7503334	1348630H1	SNP00132574	21	275	T	T	C	H80	n/a	n/a	n/a	n/a
98	7503334	1473085H1	SNP00073583	127	1100	C	C	T	H355	n/d	n/a	n/a	n/a
98	7503334	1845334H1	SNP00115181	199	1839	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7503334	1879703H1	SNP00115181	245	1841	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7503334	1979634H1	SNP00028947	57	1548	A	A	G	noncoding	n/a	n/a	n/a	n/a
98	7503334	2675044H1	SNP00070789	140	176	C	C	A	T47	0.59	n/a	n/a	n/a
98	7503334	2895120H1	SNP00028945	207	778	G	G	A	S248	n/a	n/a	n/a	n/a
98	7503334	2957969H1	SNP00028945	141	785	G	G	A	P250	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
98	7503334	3012110H1	SNP00115181	216	1840	G	G	A	noncoding	n/a	n/a	n/a	n/a
98	7503334	3533809H1	SNP00070789	245	175	C	C	A	T47	0.59	n/a	n/a	n/a
98	7503334	3588879H1	SNP00070789	91	173	C	C	A	A46	0.59	n/a	n/a	n/a
98	7503334	3588879H1	SNP00132574	191	273	T	T	C	Y80	n/a	n/a	n/a	n/a
98	7503334	3815151H1	SNP00028945	190	784	G	G	A	R250	n/a	n/a	n/a	n/a
98	7503334	3964722H1	SNP00073583	99	1097	C	C	T	Y354	n/d	n/a	n/a	n/a
98	7503334	4631942H1	SNP00005932	5	1485	G	G	C	D484	0.58	n/a	n/a	n/a
98	7503334	4744556H1	SNP00028945	212	783	G	G	A	A250	n/a	n/a	n/a	n/a
98	7503334	4823852H1	SNP00028945	136	782	G	G	A	M249	n/a	n/a	n/a	n/a
98	7503334	5018292H1	SNP00028946	66	1120	T	C	T	I362	0.79	n/a	n/a	n/a
98	7503334	6144641H1	SNP00028946	144	1121	T	C	T	S362	0.79	n/a	n/a	n/a
98	7503334	6781329H1	SNP00005932	528	1488	G	G	C	V485	0.58	n/a	n/a	n/a
98	7503334	7037030H1	SNP00073583	180	1099	T	C	T	L355	n/d	n/a	n/a	n/a
99	7503341	1993669H1	SNP00108950	209	554	C	C	T	V179	n/a	n/a	n/a	n/a
99	7503341	1993669H1	SNP00108951	220	565	A	A	G	D183	n/a	n/a	n/a	n/a
99	7503341	3421864H1	SNP00108950	214	553	C	C	T	A179	n/a	n/a	n/a	n/a
99	7503341	3421864H1	SNP00108951	225	564	A	A	G	N183	n/a	n/a	n/a	n/a
99	7503341	4842348H1	SNP00108950	152	485	C	C	T	N156	n/a	n/a	n/a	n/a
99	7503341	4842348H1	SNP00108951	163	496	A	A	G	D160	n/a	n/a	n/a	n/a
100	7509936	1879606H1	SNP00036752	249	2445	T	T	C	G771	n/a	n/a	n/a	n/a
100	7509936	2722769H1	SNP00107580	65	2652	A	C	A	A840	n/a	n/a	n/a	n/a
100	7509936	3092557H1	SNP00036752	127	2442	T	T	C	D770	n/a	n/a	n/a	n/a
100	7509936	3127147H1	SNP00036753	249	3422	A	G	A	noncoding	n/d	0.47	0.34	0.62
100	7509936	4060730H1	SNP00036753	236	3421	G	G	A	noncoding	n/d	0.47	0.34	0.62
102	7510010	1923362H1	SNP00115459	168	2507	C	C	A	noncoding	n/d	n/a	n/a	n/a
102	7510010	6780166J1	SNP00095379	467	608	C	T	C	T78	n/d	n/d	n/a	n/a
102	7510010	7352479H1	SNP00149098	259	2165	C	C	A	H597	n/a	n/a	n/a	n/a
102	7510010	8176661H1	SNP00149098	620	2264	A	C	A	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
104	7510398	1471480H1	SNP00063744	129	1079	T	T	C	S337	n/a	n/a	n/a	n/a
106	7510044	1339846H1	SNP00038040	159	1225	G	G	A	E408	n/a	n/a	n/a	n/a
106	7510044	3438171H1	SNP00038040	234	1224	G	G	A	G408	n/a	n/a	n/a	n/a
106	7510044	4301001H1	SNP00135009	255	2250	C	C	T	noncoding	n/a	n/a	n/a	n/a
106	7510044	4574128H1	SNP00038040	35	1223	G	G	A	E408	n/a	n/a	n/a	n/a
106	7510044	6735192H1	SNP00135009	129	2248	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	1394170H1	SNP00016655	172	767	T	T	G	noncoding	n/a	n/a	n/a	n/a
107	7504509	1494530H1	SNP00023734	139	277	A	A	C	G44	n/a	n/a	n/a	n/a
107	7504509	1580276H1	SNP00016656	32	1233	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	1628406H1	SNP00016654	75	584	G	G	A	noncoding	n/a	n/a	n/a	n/a
107	7504509	1628407H1	SNP00016654	85	594	G	G	A	noncoding	n/a	n/a	n/a	n/a
107	7504509	1802234H1	SNP00132864	55	991	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	1904187H1	SNP00016654	108	592	G	G	A	noncoding	n/a	n/a	n/a	n/a
107	7504509	1929619H1	SNP00016656	4	1225	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	2122009H1	SNP00016654	43	593	G	G	A	noncoding	n/a	n/a	n/a	n/a
107	7504509	2238830H1	SNP00016653	163	202	C	C	T	A19	n/a	n/a	n/a	n/a
107	7504509	2653058H1	SNP00016656	105	1232	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	2739883H1	SNP00023734	146	263	A	A	C	T40	n/a	n/a	n/a	n/a
107	7504509	2831773H1	SNP00016653	153	203	C	C	T	R20	n/a	n/a	n/a	n/a
107	7504509	2852869H1	SNP00016653	161	170	C	C	T	R9	n/a	n/a	n/a	n/a
107	7504509	2907719H1	SNP00016653	95	201	C	C	T	A19	n/a	n/a	n/a	n/a
107	7504509	3190188H1	SNP00065746	225	1269	A	A	G	noncoding	n/a	n/a	n/a	n/a
107	7504509	3285204H1	SNP00016654	244	589	G	G	A	noncoding	n/a	n/a	n/a	n/a
107	7504509	3339555H1	SNP00016653	167	199	C	C	T	N18	n/a	n/a	n/a	n/a
107	7504509	3697331H1	SNP00016655	109	765	T	T	G	noncoding	n/a	n/a	n/a	n/a
107	7504509	3926122H1	SNP00132864	152	978	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	3927459H1	SNP00132864	149	976	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	3927494H1	SNP00132864	161	986	C	C	T	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
107	7504509	3976259H1	SNP00016654	151	576	G	G	A	noncoding	n/a	n/a	n/a	n/a
107	7504509	4112161H1	SNP00016653	170	200	C	C	T	P19	n/a	n/a	n/a	n/a
107	7504509	4353902H1	SNP00016656	137	1229	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	4569892H1	SNP00065746	224	1267	A	A	G	noncoding	n/a	n/a	n/a	n/a
107	7504509	4595542H1	SNP00016654	172	587	G	G	A	noncoding	n/a	n/a	n/a	n/a
107	7504509	4648842H1	SNP00132864	50	990	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	4648886H1	SNP00132864	50	989	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	4726603H1	SNP00016656	182	1231	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	4884130H1	SNP00132864	112	974	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	5055962H2	SNP00016656	72	1228	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	5692410H1	SNP00016654	135	590	G	G	A	noncoding	n/a	n/a	n/a	n/a
107	7504509	570539H1	SNP00016656	241	1234	C	C	T	noncoding	n/a	n/a	n/a	n/a
107	7504509	6064173H1	SNP00016655	120	713	T	T	G	noncoding	n/a	n/a	n/a	n/a
107	7504509	6411454H1	SNP00065746	363	1259	A	A	G	noncoding	n/a	n/a	n/a	n/a
107	7504509	836712H1	SNP00132864	35	979	C	C	T	noncoding	n/a	n/a	n/a	n/a
108	7506825	1453355H1	SNP00016648	5	636	A	A	G	noncoding	n/a	n/a	n/a	n/a
108	7506825	167782H1	SNP00016648	45	631	A	A	G	noncoding	n/a	n/a	n/a	n/a
108	7506825	1954487H1	SNP00016648	92	635	A	A	G	noncoding	n/a	n/a	n/a	n/a
108	7506825	2209582H1	SNP00103712	49	156	G	G	A	noncoding	n/a	n/a	n/a	n/a
108	7506825	2472745H1	SNP00103712	84	155	G	G	A	noncoding	n/a	n/a	n/a	n/a
108	7506825	2516540H1	SNP00103712	48	153	G	G	A	noncoding	n/a	n/a	n/a	n/a
108	7506825	4331092H1	SNP00016648	195	633	A	A	G	noncoding	n/a	n/a	n/a	n/a
109	7506828	1453355H1	SNP00016648	5	701	A	A	G	N179	n/a	n/a	n/a	n/a
109	7506828	167782H1	SNP00016648	45	696	A	A	G	I177	n/a	n/a	n/a	n/a
109	7506828	1954487H1	SNP00016648	92	700	A	A	G	N179	n/a	n/a	n/a	n/a
109	7506828	2209582H1	SNP00103712	49	156	G	G	A	noncoding	n/a	n/a	n/a	n/a
109	7506828	2472745H1	SNP00103712	84	155	G	G	A	noncoding	n/a	n/a	n/a	n/a
109	7506828	2516540H1	SNP00103712	48	153	G	G	A	noncoding	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
109	7506828	4331092H1	SNP00016648	195	698	A	A	G	D178	n/a	n/a	n/a	n/a
109	7506828	4337014H1	SNP00103712	48	152	G	G	A	noncoding	n/a	n/a	n/a	n/a
110	7506831	1453355H1	SNP00016648	5	596	A	A	G	N144	n/a	n/a	n/a	n/a
110	7506831	167782H1	SNP00016648	45	591	A	A	G	I142	n/a	n/a	n/a	n/a
110	7506831	1954487H1	SNP00016648	92	595	A	A	G	N144	n/a	n/a	n/a	n/a
110	7506831	2209582H1	SNP00103712	49	156	G	G	A	noncoding	n/a	n/a	n/a	n/a
110	7506831	2472745H1	SNP00103712	84	155	G	G	A	noncoding	n/a	n/a	n/a	n/a
110	7506831	2516540H1	SNP00103712	48	153	G	G	A	noncoding	n/a	n/a	n/a	n/a
110	7506831	4331092H1	SNP00016648	195	593	A	A	G	D143	n/a	n/a	n/a	n/a
110	7506831	4337014H1	SNP00103712	48	152	G	G	A	noncoding	n/a	n/a	n/a	n/a
112	7510232	4541487H1	SNP00027989	33	1401	A	G	A	noncoding	n/a	n/a	n/a	n/a
112	7510232	6568762H1	SNP00027989	125	1403	G	G	A	noncoding	n/a	n/a	n/a	n/a
113	7510233	4541487H1	SNP00027989	33	1310	A	G	A	K352	n/a	n/a	n/a	n/a
113	7510233	6568762H1	SNP00027989	125	1312	G	G	A	K352	n/a	n/a	n/a	n/a
114	7510304	1796202H1	SNP00074911	132	1278	T	T	G	I386	n/d	n/d	n/d	n/d
114	7510304	1796202H1	SNP00074912	139	1285	T	T	G	W389	n/d	n/d	n/d	n/d
114	7510304	2514544H1	SNP00074911	307	1280	T	T	G	V387	n/d	n/d	n/d	n/d
114	7510304	2514544H1	SNP00074912	314	1287	T	T	G	C389	n/d	n/d	n/d	n/d
114	7510304	2514928H1	SNP00074910	147	874	C	C	T	H252	n/d	n/d	n/d	n/d
114	7510304	3962448H1	SNP00074911	16	1277	T	T	G	I386	n/d	n/d	n/d	n/d
114	7510304	4108587H1	SNP00074910	50	861	C	C	T	H247	n/d	n/d	n/d	n/d
114	7510304	4124170H1	SNP00074910	49	870	C	C	T	V250	n/d	n/d	n/d	n/d
114	7510304	4266042H1	SNP00039704	62	547	C	C	T	L143	n/a	n/a	n/a	n/a
114	7510304	4417138H1	SNP00074911	256	1279	T	T	G	L387	n/d	n/d	n/d	n/d
114	7510304	4594529H1	SNP00074911	125	1276	T	T	G	F386	n/d	n/d	n/d	n/d
114	7510304	4594529H1	SNP00074912	132	1283	T	T	G	I388	n/d	n/d	n/d	n/d
114	7510304	4647470H1	SNP00074910	23	872	C	C	T	S251	n/d	n/d	n/d	n/d
114	7510304	4649690H1	SNP00039704	53	546	C	C	T	S142	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
114	7510304	4797503H1	SNP00074912	145	1284	T	T	G	S388	n/d	n/d	n/d	n/d
114	7510304	4984847H1	SNP00074910	49	863	C	C	T	A248	n/d	n/d	n/d	n/d
114	7510304	5037989H1	SNP00074911	98	1117	T	T	G	F333	n/d	n/d	n/d	n/d
114	7510304	5037989H1	SNP00074912	105	1124	T	T	G	I335	n/d	n/d	n/d	n/d
114	7510304	5527268H1	SNP00074910	80	875	C	C	T	P252	n/d	n/d	n/d	n/d
114	7510304	5810447H1	SNP00074910	53	871	C	C	T	H251	n/d	n/d	n/d	n/d
114	7510304	5974372H1	SNP00148014	155	628	C	C	T	L170	n/a	n/a	n/a	n/a
114	7510304	6735594H1	SNP00039704	318	507	C	C	T	F129	n/a	n/a	n/a	n/a
114	7510304	6735594H1	SNP00148014	399	585	C	C	T	G155	n/a	n/a	n/a	n/a
114	7510304	7025319H1	SNP00040492	108	1396	G	G	C	V426	n/d	n/a	n/a	n/a
114	7510304	758342H1	SNP00074911	100	1274	T	T	G	V385	n/d	n/d	n/d	n/d
115	7510461	1348630H1	SNP00132574	21	345	T	T	C	H108	n/a	n/a	n/a	n/a
115	7510461	1473085H1	SNP00073583	127	1543	C	C	T	noncoding	n/d	n/a	n/a	n/a
115	7510461	1845334H1	SNP00115181	199	2282	G	G	A	noncoding	n/a	n/a	n/a	n/a
115	7510461	1879703H1	SNP00115181	245	2284	G	G	A	noncoding	n/a	n/a	n/a	n/a
115	7510461	1979634H1	SNP00028947	57	1991	A	A	G	noncoding	n/a	n/a	n/a	n/a
115	7510461	2675044H1	SNP00070789	140	246	C	C	A	T75	0.59	n/a	n/a	n/a
115	7510461	2895120H1	SNP00028945	207	848	G	G	A	S276	n/a	n/a	n/a	n/a
115	7510461	2957969H1	SNP00028945	141	855	G	G	A	P278	n/a	n/a	n/a	n/a
115	7510461	3012110H1	SNP00115181	216	2283	G	G	A	noncoding	n/a	n/a	n/a	n/a
115	7510461	3533809H1	SNP00070789	245	245	C	C	A	T75	0.59	n/a	n/a	n/a
115	7510461	3588879H1	SNP00070789	91	243	C	C	A	A74	0.59	n/a	n/a	n/a
115	7510461	3588879H1	SNP00132574	191	343	T	T	C	Y108	n/a	n/a	n/a	n/a
115	7510461	3815151H1	SNP00028945	190	854	G	G	A	R278	n/a	n/a	n/a	n/a
115	7510461	3964722H1	SNP00073583	99	1540	C	C	T	noncoding	n/d	n/a	n/a	n/a
115	7510461	4631942H1	SNP00005932	5	1928	G	G	C	noncoding	0.58	n/a	n/a	n/a
115	7510461	4744556H1	SNP00028945	212	853	G	G	A	A278	n/a	n/a	n/a	n/a
115	7510461	4823852H1	SNP00028945	136	852	G	G	A	M277	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
115	7510461	5018292H1	SNP00028946	66	1563	T	C	T	noncoding	0.79	n/a	n/a	n/a
115	7510461	6781329H1	SNP00005932	528	1931	G	G	C	noncoding	0.58	n/a	n/a	n/a
115	7510461	7037030H1	SNP00073583	180	1542	T	C	T	noncoding	n/d	n/a	n/a	n/a
115	7510461	7354101H1	SNP00028946	9	1564	C	C	T	noncoding	0.79	n/a	n/a	n/a
115	7510461	8504647H1	SNP00028946	197	1194	C	C	T	P391	0.79	n/a	n/a	n/a
116	7510392	1313357H1	SNP00061075	9	241	C	C	T	noncoding	n/a	n/a	n/a	n/a
116	7510392	3350442H1	SNP00061075	4	240	C	C	T	noncoding	n/a	n/a	n/a	n/a
116	7510392	4713619H1	SNP00061075	10	239	C	C	T	noncoding	n/a	n/a	n/a	n/a
116	7510392	6054452H1	SNP00061075	206	224	C	C	T	noncoding	n/a	n/a	n/a	n/a

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58,

b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:6-7, SEQ ID NO:10-12, SEQ ID NO:15, SEQ ID NO:17-18, SEQ ID NO:21, SEQ ID NO:23-25, SEQ ID NO:27, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:41-42, SEQ ID NO:45, SEQ ID NO:50, and SEQ ID NO:57,

c) a polypeptide comprising a naturally occurring amino acid sequence at least 95% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:5, SEQ ID NO:19, and SEQ ID NO:58,

d) a polypeptide comprising a naturally occurring amino acid sequence at least 93% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:8 and SEQ ID NO:32,

e) a polypeptide comprising a naturally occurring amino acid sequence at least 99% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:13, SEQ ID NO:26, and SEQ ID NO:36,

f) a polypeptide comprising a naturally occurring amino acid sequence at least 91% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:16, SEQ ID NO: 47, and SEQ ID NO:56,

g) a polypeptide comprising a naturally occurring amino acid sequence at least 92% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:49 and SEQ ID NO:56,

h) a polypeptide comprising a naturally occurring amino acid sequence at least 96% identical to the amino acid sequence of SEQ ID NO:44,

i) a polypeptide comprising a naturally occurring amino acid sequence at least 97% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:40 and SEQ ID NO:48,

j) a polypeptide comprising a naturally occurring amino acid sequence at least 98% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:54-55,

k) a polypeptide consisting essentially of a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:4, SEQ ID NO:9, SEQ ID NO:14, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:28-31, SEQ ID NO:34, SEQ ID NO:38-39, SEQ ID NO:43, SEQ ID NO:46, and SEQ ID NO:51-53,

l) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, and

m) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-58.

2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method of producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-58.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:59-116,

b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:61-66, SEQ ID NO:68-70, SEQ ID NO:73-75, SEQ ID NO:79, SEQ ID NO:81-82, SEQ ID NO:84, SEQ ID NO:88, SEQ ID NO:91-93, SEQ ID NO:95, SEQ ID NO:98, SEQ ID NO:111, and SEQ ID NO:115-116,

c) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 98% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:59 and SEQ ID NO:101-102,

d) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 95% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:60 and SEQ ID NO:103,

e) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 99% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:67, SEQ ID NO:72, SEQ ID NO:96, and SEQ ID NO:99-100,

f) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 97% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:87, SEQ ID NO:97, and SEQ ID NO:105,

g) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 91% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:104 and SEQ ID NO:113,

h) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 92% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:109-110,

i) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 93% identical to the polynucleotide sequence of SEQ ID NO:108,

j) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 94% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:107 and SEQ ID NO:114,

k) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 96% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:106, SEQ ID NO:86, SEQ ID NO:94, and SEQ ID NO:112,

l) a polynucleotide consisting essentially of a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:76-78,

m) a polynucleotide complementary to a polynucleotide of a),

n) a polynucleotide complementary to a polynucleotide of b),

o) a polynucleotide complementary to a polynucleotide of c),

p) a polynucleotide complementary to a polynucleotide of d),

q) a polynucleotide complementary to a polynucleotide of e),

r) a polynucleotide complementary to a polynucleotide of f),

s) a polynucleotide complementary to a polynucleotide of g),

t) a polynucleotide complementary to a polynucleotide of h),

u) a polynucleotide complementary to a polynucleotide of i),

v) a polynucleotide complementary to a polynucleotide of j),

w) a polynucleotide complementary to a polynucleotide of k),

x) a polynucleotide complementary to a polynucleotide of l), and

y) an RNA equivalent of a)-x).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-58.

19. A method for treating a disease or condition associated with decreased expression of functional PMMM, comprising administering to a patient in need of such treatment the composition of claim 17.

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of functional PMMM, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

25. A method for treating a disease or condition associated with overexpression of functional PMMM, comprising administering to a patient in need of such treatment a composition of claim 24.

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

a) treating a biological sample containing nucleic acids with the test compound,

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,

c) quantifying the amount of hybridization complex, and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

30. A method for a diagnostic test for a condition or disease associated with the expression of PMMM in a biological sample, the method comprising:

a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

31. The antibody of claim 11, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

33. A method of diagnosing a condition or disease associated with the expression of PMMM in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, further comprising a label.

35. A method of diagnosing a condition or disease associated with the expression of PMMM in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibodies from the animal, and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58.

37. A polyclonal antibody produced by a method of claim 36.

38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-58, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibody producing cells from the animal,

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,

d) culturing the hybridoma cells, and

e) isolating from the culture monoclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58.

40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58 in a sample, the method comprising:

a) incubating the antibody of claim 11 with the sample under conditions to allow specific binding of the antibody and the polypeptide, and
b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58 in the sample.

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58 from a sample, the method comprising:

a) incubating the antibody of claim 11 with the sample under conditions to allow specific binding of the antibody and the polypeptide, and
b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-58.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

a) labeling the polynucleotides of the sample,

b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.

89. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.

90. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.

91. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.

92. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.

93. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

94. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.

95. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.

96. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.

97. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.

98. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:43.

99. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:44.

100. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:45.

101. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:46.

102. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:47.

103. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:48.

104. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:49.

105. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:50.

106. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:51.

107. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:52.

108. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:53.

109. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:54.

110. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:55.

111. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:56.

112. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:57.

113. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:58.

114. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:59.

115. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:60.

116. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:61.

117. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:62.

118. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:63.

119. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:64.

120. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:65.

121. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:66.

122. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:67.

123. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:68.

124. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:69.

125. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:70.

126. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:71.

127. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:72.

128. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:73.

129. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:74.

130. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:75.

131. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:76.

132. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:77.

133. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:78.

134. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:79.

135. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:80.

136. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:81.

137. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:82.

138. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:83.

139. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:84.

140. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:85.

141. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:86.

142. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:87.

143. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:88.

144. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:89.

145. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:90.

146. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:91.

147. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:92.

148. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:93.

149. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:94.

150. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:95.

151. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:96.

152. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:97.

153. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:98.

154. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:99.

155. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:100.

156. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:101.

157. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:102.

158. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:103.

159. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:104.

160. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:105.

161. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:106.

162. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:107.

163. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:108.

164. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:109.

165. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:110.

166. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:111.

167. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:112.

168. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:113.

169. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:114.

170. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:115.

171. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:116.

<110> INCYTE GENOMICS, INC.; HAFALIA, April J.A.;
LI, Joana X.; GORVAD, Ann E.;
CHAWLA, Narinder K.; SPRAGUE, William W.;
LEE, Soo Yeun; CHANG, Hsin-Ru;
ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi;
KHARE, Reena; EMERLING, Brooke M.;
KABLE, Amy E.; TANG, Y. Tom;
YUE, Henry; GIETZEN, Kimberly J.;
LEE, Sally; SWARNAKAR, Anita;
BAUGHN, Mariah R.; WILSON, Amy D.;
JIN, Pei; CHIEN, David;
HAWKINS, Phillip R.; JIANG, Xin;
JACKSON, Alan A.; BHATIA, Umesh;
BURRILL, John D.; BLAKE, Julie J.;
HO, Anne; ZHENG, Wenjin;
ISON, Craig H.; MARQUIS, Joseph P.;
TRAN, Uyen K.; LAL, Preeti G.;
WARREN, Bridget A.; XU, Yuming;
HONCHELL, Cynthia D.; BECHA, Shanya D.;
LEHR-MASON, Patricia M.

<120> PROTEIN MODIFICATION AND MAINTENANCE MOLECULES

<130> PF-1362 PCT

<140> To Be Assigned
<141> Herewith

<150> US 60/351,928
<151> 2002-01-25

<150> US 60/359,903
<151> 2002-02-25

<150> US 60/366,837
<151> 2002-03-21

<160> 116
<170> PERL Program

<210> 1
<211> 321
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7313196CD1

<400> 1
```
Met Asn Trp His Leu Pro Leu Phe Leu Leu Ala Ser Val Thr Leu
 1               5                  10                  15
Pro Ser Ile Cys Ser His Phe Asn Pro Leu Ser Leu Glu Glu Leu
                20                  25                  30
Gly Ser Asn Thr Gly Ile Gln Val Phe Asn Gln Ile Val Lys Ser
                35                  40                  45
Arg Pro His Asp Asn Ile Val Ile Ser Pro His Gly Ile Ala Ser
                50                  55                  60
Val Leu Gly Met Leu Gln Leu Gly Ala Asp Gly Arg Thr Lys Lys
                65                  70                  75
Gln Leu Ala Met Val Met Arg Tyr Gly Val Asn Asp Met Ile Asp
                80                  85                  90
Asn Leu Leu Ser Pro Asp Leu Ile Asp Gly Val Leu Thr Arg Leu
```

```
                95                      100                     105
Val Leu Val Asn Ala Val Tyr Phe Lys Gly Leu Trp Lys Ser Arg
                110                     115                     120
Phe Gln Pro Glu Asn Thr Lys Lys Arg Thr Phe Val Ala Ala Asp
                125                     130                     135
Gly Lys Ser Tyr Gln Val Pro Met Leu Ala Gln Leu Ser Val Phe
                140                     145                     150
Arg Cys Gly Ser Thr Ser Ala Pro Asn Asp Leu Trp Tyr Asn Phe
                155                     160                     165
Ile Glu Leu Pro Tyr His Gly Glu Ser Ile Ser Met Leu Ile Ala
                170                     175                     180
Leu Pro Thr Glu Ser Ser Thr Pro Leu Ser Ala Ile Ile Pro His
                185                     190                     195
Ile Ser Thr Lys Thr Ile Asp Ser Trp Met Ser Ile Met Val Pro
                200                     205                     210
Lys Arg Val Gln Val Ile Leu Pro Lys Phe Thr Ala Val Ala Gln
                215                     220                     225
Thr Asp Leu Lys Glu Pro Leu Lys Val Leu Gly Ile Thr Asp Met
                230                     235                     240
Phe Asp Ser Ser Lys Ala Asn Phe Ala Lys Ile Thr Arg Ser Glu
                245                     250                     255
Asn Leu His Val Ser His Ile Leu Gln Lys Ala Lys Ile Glu Val
                260                     265                     270
Ser Glu Asp Gly Thr Lys Ala Ser Ala Ala Thr Thr Ala Ile Leu
                275                     280                     285
Ile Ala Arg Ser Ser Pro Pro Trp Phe Ile Val Asp Arg Pro Phe
                290                     295                     300
Leu Phe Phe Ile Arg His Asn Pro Thr Gly Ala Val Leu Phe Met
                305                     310                     315
Gly Gln Ile Asn Lys Pro
                320
```

<210> 2
<211> 1309
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6465289CD1

<400> 2

```
Met Val Pro Gly Glu Glu Asn Gln Leu Val Pro Lys Glu Ile Glu
 1               5                      10                      15
Asn Ala Ala Glu Glu Pro Arg Val Leu Cys Ile Ile Gln Asp Thr
                20                      25                      30
Thr Asn Ser Lys Thr Val Asn Glu Arg Ile Thr Leu Asn Leu Pro
                35                      40                      45
Ala Ser Thr Pro Val Arg Lys Leu Phe Glu Asp Val Ala Asn Lys
                50                      55                      60
Val Gly Tyr Ile Asn Gly Thr Phe Asp Leu Val Trp Gly Asn Gly
                65                      70                      75
Ile Asn Thr Ala Asp Met Ala Pro Leu Asp His Thr Ser Asp Lys
                80                      85                      90
Ser Leu Leu Asp Ala Asn Phe Glu Pro Gly Lys Lys Asn Phe Leu
                95                      100                     105
His Leu Thr Asp Lys Asp Gly Glu Gln Pro Gln Ile Leu Leu Glu
                110                     115                     120
Asp Ser Ser Ala Gly Glu Asp Ser Val His Asp Arg Phe Ile Gly
                125                     130                     135
Pro Leu Pro Arg Glu Gly Ser Val Gly Ser Thr Ser Asp Tyr Val
                140                     145                     150
Ser Gln Ser Tyr Ser Tyr Ser Ser Ile Leu Asn Lys Ser Glu Thr
```

155 160 165
Gly Tyr Val Gly Leu Val Asn Gln Ala Met Thr Cys Tyr Leu Asn
170 175 180
Ser Leu Leu Gln Thr Leu Phe Met Thr Pro Glu Phe Arg Asn Ala
185 190 195
Leu Tyr Lys Trp Glu Phe Glu Glu Ser Glu Glu Asp Pro Val Thr
200 205 210
Ser Ile Pro Tyr Gln Leu Gln Arg Leu Phe Val Leu Leu Gln Thr
215 220 225
Ser Lys Lys Arg Ala Ile Glu Thr Thr Asp Val Thr Arg Ser Phe
230 235 240
Gly Trp Asp Ser Ser Glu Ala Trp Gln Gln His Asp Val Gln Glu
245 250 255
Leu Cys Arg Val Met Phe Asp Ala Leu Glu Gln Lys Trp Lys Gln
260 265 270
Thr Glu Gln Ala Asp Leu Ile Asn Glu Leu Tyr Gln Gly Lys Leu
275 280 285
Lys Asp Tyr Val Arg Cys Leu Glu Cys Gly Tyr Glu Gly Trp Arg
290 295 300
Ile Asp Thr Tyr Leu Asp Ile Pro Leu Val Ile Arg Pro Tyr Gly
305 310 315
Ser Ser Gln Ala Phe Ala Ser Val Glu Glu Ala Leu His Ala Phe
320 325 330
Ile Gln Pro Glu Ile Leu Asp Gly Pro Asn Gln Tyr Phe Cys Glu
335 340 345
Arg Cys Lys Lys Lys Cys Asp Ala Arg Lys Gly Leu Arg Phe Leu
350 355 360
His Phe Pro Tyr Leu Leu Thr Leu Gln Leu Lys Arg Phe Asp Phe
365 370 375
Asp Tyr Thr Thr Met His Arg Ile Lys Leu Asn Asp Arg Met Thr
380 385 390
Phe Pro Glu Glu Leu Asp Met Ser Thr Phe Ile Asp Val Glu Asp
395 400 405
Glu Lys Ser Pro Gln Thr Glu Ser Cys Thr Asp Ser Gly Ala Glu
410 415 420
Asn Glu Gly Ser Cys His Ser Asp Gln Met Ser Asn Asp Phe Ser
425 430 435
Asn Asp Asp Gly Val Asp Glu Gly Ile Cys Leu Glu Thr Asn Ser
440 445 450
Gly Thr Glu Lys Ile Ser Lys Ser Gly Leu Glu Lys Asn Ser Leu
455 460 465
Ile Tyr Glu Leu Phe Ser Val Met Val His Ser Gly Ser Ala Ala
470 475 480
Gly Gly His Tyr Tyr Ala Cys Ile Lys Ser Phe Ser Asp Glu Gln
485 490 495
Trp Tyr Ser Phe Asn Asp Gln His Val Ser Arg Ile Thr Gln Glu
500 505 510
Asp Ile Lys Lys Thr His Gly Gly Ser Ser Gly Ser Arg Gly Tyr
515 520 525
Tyr Ser Ser Ala Phe Ala Ser Ser Thr Asn Ala Tyr Met Leu Ile
530 535 540
Tyr Arg Leu Lys Asp Pro Ala Arg Asn Ala Lys Phe Leu Glu Val
545 550 555
Asp Glu Tyr Pro Glu His Ile Lys Asn Leu Val Gln Lys Glu Arg
560 565 570
Glu Leu Glu Glu Gln Glu Lys Arg Gln Arg Glu Ile Glu Arg Asn
575 580 585
Thr Cys Lys Ile Lys Leu Phe Cys Leu His Pro Thr Lys Gln Val
590 595 600
Met Met Glu Asn Lys Leu Glu Val His Lys Asp Lys Thr Leu Lys
605 610 615
Glu Ala Val Glu Met Ala Tyr Lys Met Met Asp Leu Glu Glu Val
620 625 630

Ile Pro Leu Asp Cys Cys Arg Leu Val Lys Tyr Asp Glu Phe His
635 640 645
Asp Tyr Leu Glu Arg Ser Tyr Glu Gly Glu Glu Asp Thr Pro Met
650 655 660
Gly Leu Leu Leu Gly Gly Val Lys Ser Thr Tyr Met Phe Asp Leu
665 670 675
Leu Leu Glu Thr Arg Lys Pro Asp Gln Val Phe Gln Ser Tyr Lys
680 685 690
Pro Gly Glu Val Met Val Lys Val His Val Val Asp Leu Lys Ala
695 700 705
Glu Ser Val Ala Ala Pro Ile Thr Val Arg Ala Tyr Leu Asn Gln
710 715 720
Thr Val Thr Glu Phe Lys Gln Leu Ile Ser Lys Ala Ile His Leu
725 730 735
Pro Ala Glu Thr Met Arg Ile Val Leu Glu Arg Cys Tyr Asn Asp
740 745 750
Leu Arg Leu Leu Ser Val Ser Ser Lys Thr Leu Lys Ala Glu Gly
755 760 765
Phe Phe Arg Ser Asn Lys Val Phe Val Glu Ser Ser Glu Thr Leu
770 775 780
Asp Tyr Gln Met Ala Phe Ala Asp Ser His Leu Trp Lys Leu Leu
785 790 795
Asp Arg His Ala Asn Thr Ile Arg Leu Phe Val Leu Leu Pro Glu
800 805 810
Gln Ser Pro Val Ser Tyr Ser Lys Arg Thr Ala Tyr Gln Lys Ala
815 820 825
Gly Gly Asp Ser Gly Asn Val Asp Asp Asp Cys Glu Arg Val Lys
830 835 840
Gly Pro Val Gly Ser Leu Lys Ser Val Glu Ala Ile Leu Glu Glu
845 850 855
Ser Thr Glu Lys Leu Lys Ser Leu Ser Leu Gln Gln Gln Gln Asp
860 865 870
Gly Asp Asn Gly Asp Ser Ser Lys Ser Thr Glu Thr Ser Asp Phe
875 880 885
Glu Asn Ile Glu Ser Pro Leu Asn Glu Arg Asp Ser Ser Ala Ser
890 895 900
Val Asp Asn Arg Glu Leu Glu Gln His Ile Gln Thr Ser Asp Pro
905 910 915
Glu Asn Phe Gln Ser Glu Glu Arg Ser Asp Ser Asp Val Asn Asn
920 925 930
Asp Arg Ser Thr Ser Ser Val Asp Ser Asp Ile Leu Ser Ser Ser
935 940 945
His Ser Ser Asp Thr Leu Cys Asn Ala Asp Asn Ala Gln Ile Pro
950 955 960
Leu Ala Asn Gly Leu Asp Ser His Ser Ile Thr Ser Ser Arg Arg
965 970 975
Thr Lys Ala Asn Glu Gly Lys Lys Glu Thr Trp Asp Thr Ala Glu
980 985 990
Glu Asp Ser Gly Thr Asp Ser Glu Tyr Asp Glu Ser Gly Lys Ser
995 1000 1005
Arg Gly Glu Met Gln Tyr Met Tyr Phe Lys Ala Glu Pro Tyr Ala
1010 1015 1020
Ala Asp Glu Gly Ser Gly Glu Gly His Lys Trp Leu Met Val His
1025 1030 1035
Val Asp Lys Arg Ile Thr Leu Ala Ala Phe Lys Gln His Leu Glu
1040 1045 1050
Pro Phe Val Gly Val Leu Ser Ser His Phe Lys Val Phe Arg Val
1055 1060 1065
Tyr Ala Ser Asn Gln Glu Phe Glu Ser Val Arg Leu Asn Glu Thr
1070 1075 1080
Leu Ser Ser Phe Ser Asp Asp Asn Lys Ile Thr Ile Arg Leu Gly
1085 1090 1095
Arg Ala Leu Lys Lys Gly Glu Tyr Arg Val Lys Val Tyr Gln Leu

1100 1105 1110
Leu Val Asn Glu Gln Glu Pro Cys Lys Phe Leu Leu Asp Ala Val
1115 1120 1125
Phe Ala Lys Gly Met Thr Val Arg Gln Ser Lys Glu Glu Leu Ile
1130 1135 1140
Pro Gln Leu Arg Glu Gln Cys Gly Leu Glu Leu Ser Ile Asp Arg
1145 1150 1155
Phe Arg Leu Arg Lys Lys Thr Trp Lys Asn Pro Gly Thr Val Phe
1160 1165 1170
Leu Asp Tyr His Ile Tyr Glu Glu Asp Ile Asn Ile Ser Ser Asn
1175 1180 1185
Trp Glu Val Phe Leu Glu Val Leu Asp Gly Val Glu Lys Met Lys
1190 1195 1200
Ser Met Ser Gln Leu Ala Val Leu Ser Arg Arg Trp Lys Pro Ser
1205 1210 1215
Glu Met Lys Leu Asp Pro Phe Gln Glu Val Val Leu Glu Ser Ser
1220 1225 1230
Ser Val Asp Glu Leu Arg Glu Lys Leu Ser Glu Ile Ser Gly Ile
1235 1240 1245
Pro Leu Asp Asp Ile Glu Phe Ala Lys Gly Arg Gly Thr Phe Pro
1250 1255 1260
Cys Asp Ile Ser Val Leu Asp Ile His Gln Asp Leu Asp Trp Asn
1265 1270 1275
Pro Lys Val Ser Thr Leu Asn Val Trp Pro Leu Tyr Ile Cys Asp
1280 1285 1290
Asp Gly Ala Val Ile Phe Tyr Arg Asp Lys Thr Glu Glu Leu Met
1295 1300 1305
Glu Ile Asp Arg

<210> 3
<211> 987
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506357CD1

<400> 3
Met Asp Ala Leu Asp Ala Ser Lys Leu Leu Asp Glu Glu Leu Tyr
1 5 10 15
Ser Arg Gln Leu Tyr Val Leu Gly Ser Pro Ala Met Gln Arg Ile
20 25 30
Gln Gly Ala Arg Val Leu Val Ser Gly Leu Gln Gly Leu Gly Ala
35 40 45
Glu Val Ala Lys Asn Leu Val Leu Met Gly Val Gly Ser Leu Thr
50 55 60
Leu His Asp Pro His Pro Thr Cys Trp Ser Asp Leu Ala Ala Gln
65 70 75
Phe Leu Leu Ser Glu Gln Asp Leu Glu Arg Ser Arg Ala Glu Ala
80 85 90
Ser Gln Glu Leu Leu Ala Gln Leu Asn Arg Ala Val Gln Val Val
95 100 105
Val His Thr Gly Asp Ile Thr Glu Asp Leu Leu Leu Asp Phe Gln
110 115 120
Val Val Val Leu Thr Ala Ala Lys Leu Glu Glu Gln Leu Lys Val
125 130 135
Gly Thr Leu Cys His Lys His Gly Val Cys Phe Leu Ala Ala Asp
140 145 150
Thr Arg Gly Leu Val Gly Gln Leu Phe Cys Asp Phe Gly Glu Asp
155 160 165
Phe Thr Val Gln Asp Pro Thr Glu Ala Glu Pro Leu Thr Ala Ala

170 175 180
Ile Gln His Ile Ser Gln Gly Ser Pro Gly Ile Leu Thr Leu Arg
185 190 195
Lys Gly Ala Asn Thr His Tyr Phe Arg Asp Gly Asp Leu Val Thr
200 205 210
Phe Ser Gly Ile Glu Gly Met Val Glu Leu Asn Asp Cys Asp Pro
215 220 225
Arg Ser Ile His Val Arg Glu Asp Gly Ser Leu Glu Ile Gly Asp
230 235 240
Thr Thr Thr Phe Ser Arg Tyr Leu Arg Gly Gly Ala Ile Thr Glu
245 250 255
Val Lys Arg Pro Lys Thr Val Arg His Lys Ser Leu Asp Thr Ala
260 265 270
Leu Leu Gln Pro His Val Val Ala Gln Ser Ser Gln Glu Val His
275 280 285
His Ala His Cys Leu His Gln Ala Phe Cys Ala Leu His Lys Phe
290 295 300
Gln His Leu His Gly Arg Pro Pro Gln Pro Trp Asp Pro Val Asp
305 310 315
Ala Glu Thr Val Val Gly Leu Ala Arg Asp Leu Glu Pro Leu Lys
320 325 330
Arg Thr Glu Glu Glu Pro Leu Glu Glu Pro Leu Asp Glu Ala Leu
335 340 345
Val Arg Thr Val Ala Leu Ser Ser Ala Gly Val Leu Ser Pro Met
350 355 360
Val Ala Met Leu Gly Ala Val Ala Ala Gln Glu Val Leu Lys Ala
365 370 375
Ile Ser Arg Lys Phe Met Pro Leu Asp Gln Trp Leu Tyr Phe Asp
380 385 390
Ala Leu Asp Cys Leu Pro Glu Asp Gly Glu Leu Leu Pro Ser Pro
395 400 405
Glu Asp Cys Ala Leu Arg Gly Ser Arg Tyr Asp Gly Gln Ile Ala
410 415 420
Val Phe Gly Ala Gly Phe Gln Glu Lys Leu Arg Arg Gln His Tyr
425 430 435
Leu Leu Val Gly Ala Gly Ala Ile Gly Cys Glu Leu Leu Lys Val
440 445 450
Phe Ala Leu Val Gly Leu Gly Ala Gly Asn Ser Gly Gly Leu Thr
455 460 465
Val Val Asp Met Asp His Ile Glu Arg Ser Asn Leu Ser Arg Gln
470 475 480
Phe Leu Phe Arg Ser Gln Asp Val Gly Arg Pro Lys Ala Glu Val
485 490 495
Ala Ala Ala Ala Ala Arg Gly Leu Asn Pro Asp Leu Gln Val Ile
500 505 510
Pro Leu Thr Tyr Pro Leu Asp Pro Thr Thr Glu His Ile Tyr Gly
515 520 525
Asp Asn Phe Phe Ser Arg Val Asp Gly Val Ala Ala Ala Leu Asp
530 535 540
Ser Phe Gln Ala Arg Arg Tyr Val Ala Ala Arg Cys Thr His Tyr
545 550 555
Leu Lys Pro Leu Leu Glu Ala Gly Thr Ser Gly Thr Trp Gly Ser
560 565 570
Ala Thr Val Phe Met Pro His Val Thr Glu Ala Tyr Arg Ala Pro
575 580 585
Ala Ser Ala Ala Ala Ser Glu Asp Ala Pro Tyr Pro Val Cys Thr
590 595 600
Val Arg Tyr Phe Pro Ser Thr Ala Glu His Thr Leu Gln Trp Ala
605 610 615
Arg His Glu Phe Glu Glu Leu Phe Arg Leu Ser Ala Glu Thr Ile
620 625 630
Asn His His Gln Gln Ala His Thr Ser Leu Ala Asp Met Asp Glu
635 640 645

```
Pro Gln Thr Leu Thr Leu Leu Lys Pro Val Leu Gly Val Leu Arg
                650                 655                 660
Val Arg Pro Gln Asn Trp Gln Asp Cys Val Ala Trp Ala Leu Gly
                665                 670                 675
His Trp Lys Leu Cys Phe His Tyr Gly Ile Lys Gln Leu Leu Arg
                680                 685                 690
His Phe Pro Pro Asn Lys Asp Thr His Leu Leu Tyr Val Leu Ala
                695                 700                 705
Ala Ala Asn Leu Tyr Ala Gln Met His Gly Leu Pro Gly Ser Gln
                710                 715                 720
Asp Trp Thr Ala Leu Arg Glu Leu Leu Lys Leu Leu Pro Gln Pro
                725                 730                 735
Asp Pro Gln Gln Met Ala Pro Ile Phe Ala Ser Asn Leu Glu Leu
                740                 745                 750
Ala Ser Ala Ser Ala Glu Phe Gly Pro Glu Gln Gln Lys Glu Leu
                755                 760                 765
Asn Lys Ala Leu Glu Val Trp Ser Val Gly Pro Pro Leu Lys Pro
                770                 775                 780
Leu Met Phe Glu Lys Asp Asp Asp Ser Asn Phe His Val Asp Phe
                785                 790                 795
Val Val Ala Ala Ala Ser Leu Arg Cys Gln Asn Tyr Gly Ile Pro
                800                 805                 810
Pro Val Asn Arg Ala Gln Ser Lys Arg Ile Val Gly Gln Ile Ile
                815                 820                 825
Pro Ala Ile Ala Thr Thr Thr Ala Ala Val Ala Gly Leu Leu Gly
                830                 835                 840
Leu Glu Leu Tyr Lys Val Val Ser Gly Pro Arg Pro Arg Ser Ala
                845                 850                 855
Phe Arg His Ser Tyr Leu His Leu Ala Glu Asn Tyr Leu Ile Arg
                860                 865                 870
Tyr Met Pro Phe Ala Pro Ala Ile Gln Thr Phe His His Leu Lys
                875                 880                 885
Trp Thr Ser Trp Asp Arg Leu Lys Val Pro Ala Gly Gln Pro Glu
                890                 895                 900
Arg Thr Leu Glu Ser Leu Leu Ala His Leu Gln Glu Gln His Gly
                905                 910                 915
Leu Arg Val Arg Ile Leu Leu His Gly Ser Ala Leu Leu Tyr Ala
                920                 925                 930
Ala Gly Trp Ser Pro Glu Lys Gln Ala Gln His Leu Pro Leu Arg
                935                 940                 945
Val Thr Glu Leu Val Gln Gln Leu Thr Gly Gln Ala Pro Ala Pro
                950                 955                 960
Gly Gln Arg Val Leu Val Leu Glu Leu Ser Cys Glu Gly Asp Asp
                965                 970                 975
Glu Asp Thr Ala Phe Pro Pro Leu His Tyr Glu Leu
                980                 985

&lt;210&gt; 4
&lt;211&gt; 227
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;220&gt;
&lt;221&gt; misc_feature
&lt;223&gt; Incyte ID No: 6878857CD1

&lt;400&gt; 4
Met Asn Leu Leu Leu Ile Leu Thr Phe Val Ala Ala Ala Val Ala
 1               5                  10                  15
Ala Pro Phe Asp Asp Asp Asp Lys Ile Val Gly Gly Tyr Ile Cys
                20                  25                  30
Glu Glu Asn Ser Val Pro Tyr Gln Val Ser Leu Asn Ser Gly Tyr
                35                  40                  45
```

Lys Ser Arg Ile Gln Val Arg Leu Gly Glu His Asn Ile Glu Val
50 55 60
Leu Glu Gly Asn Glu Gln Phe Ile Asn Ala Ala Lys Ile Ile Arg
65 70 75
His Pro Lys Tyr Asn Ser Arg Thr Leu Asp Asn Asp Ile Leu Leu
80 85 90
Ile Lys Leu Ser Ser Pro Ala Val Ile Asn Ser Arg Val Ser Ala
95 100 105
Ile Ser Leu Pro Thr Ala Pro Pro Ala Ala Gly Thr Glu Ser Leu
110 115 120
Ile Ser Gly Trp Gly Asn Thr Leu Ser Ser Gly Ala Asp Tyr Pro
125 130 135
Asp Glu Leu Gln Cys Leu Asp Ala Pro Val Leu Ser Gln Ala Glu
140 145 150
Cys Glu Ala Ser Tyr Pro Gly Lys Ile Thr Asn Asn Met Phe Cys
155 160 165
Val Gly Phe Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser
170 175 180
Gly Gly Pro Val Val Ser Asn Gly Glu Leu Gln Gly Ile Val Ser
185 190 195
Trp Gly Tyr Gly Cys Ala Gln Lys Asn Arg Pro Gly Val Tyr Thr
200 205 210
Lys Val Tyr Asn Tyr Val Asp Trp Ile Lys Asp Thr Ile Ala Ala
215 220 225
Asn Ser

<210> 5
<211> 727
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506021CD1

<400> 5
Met Lys Lys Gln Arg Lys Ile Leu Trp Arg Lys Gly Ile His Leu
1 5 10 15
Ala Phe Ser Glu Lys Trp Asn Thr Gly Phe Gly Gly Phe Lys Lys
20 25 30
Phe Tyr Phe His Gln His Leu Cys Ile Leu Lys Ala Lys Leu Gly
35 40 45
Arg Pro Val Thr Trp Asn Arg Gln Leu Arg His Phe Gln Gly Arg
50 55 60
Lys Lys Ala Leu Gln Ile Gln Lys Thr Trp Ile Lys Asp Glu Pro
65 70 75
Leu Cys Ala Lys Thr Lys Phe Asn Val Ala Thr Gln Asn Val Ser
80 85 90
Thr Leu Ser Ser Lys Val Lys Arg Lys Asp Ala Lys His Phe Ile
95 100 105
Ser Ser Ser Lys Thr Leu Leu Arg Leu Gln Ala Glu Lys Leu Leu
110 115 120
Ser Ser Ala Lys Asn Ser Asp His Glu Tyr Cys Arg Glu Lys Asn
125 130 135
Leu Leu Lys Ala Val Thr Asp Phe Pro Ser Asn Ser Ala Leu Gly
140 145 150
Gln Ala Asn Gly His Arg Pro Arg Thr Asp Pro Gln Pro Ser Asp
155 160 165
Phe Pro Met Lys Phe Asn Gly Glu Ser Gln Ser Pro Val Glu Ser
170 175 180
Gly Thr Ile Val Val Thr Leu Asn Asn His Lys Arg Lys Gly Phe
185 190 195

Cys Tyr Arg Cys Cys Gln Gly Pro Glu His His Arg Asn Gly Gly
200 205 210
Pro Leu Ile Pro Lys Lys Phe Gln Leu Asn Gln His Arg Arg Ile
215 220 225
Lys Leu Ser Pro Leu Met Met Tyr Glu Lys Leu Ser Met Ile Arg
230 235 240
Phe Arg Tyr Arg Ile Leu Arg Ser Gln His Phe Arg Thr Lys Ser
245 250 255
Lys Val Cys Lys Leu Arg Lys Ala Gln Arg Ser Trp Val Gln Lys
260 265 270
Val Thr Gly Asp His Gln Glu Thr Arg Arg Glu Asn Gly Glu Gly
275 280 285
Gly Ser Cys Ser Pro Phe Pro Ser Pro Glu Pro Lys Asp Pro Ser
290 295 300
Cys Arg His Gln Pro Tyr Phe Pro Asp Met Asp Ser Ser Ala Val
305 310 315
Val Lys Gly Thr Asn Ser His Val Pro Asp Cys His Thr Lys Gly
320 325 330
Ser Ser Phe Leu Gly Lys Glu Leu Ser Leu Asp Glu Ala Phe Pro
335 340 345
Asp Gln Gln Asn Gly Ser Ala Thr Asn Ala Trp Asp Gln Ser Ser
350 355 360
Cys Ser Ser Pro Lys Trp Glu Cys Thr Glu Leu Ile His Asp Ile
365 370 375
Pro Leu Pro Glu His Arg Ser Asn Thr Met Phe Ile Ser Glu Thr
380 385 390
Glu Arg Glu Ile Met Thr Leu Gly Gln Glu Asn Gln Thr Ser Ser
395 400 405
Val Ser Asp Asp Arg Val Lys Leu Ser Val Ser Gly Ala Asp Thr
410 415 420
Ser Val Ser Ser Val Asp Gly Pro Val Ser Gln Lys Ala Val Gln
425 430 435
Asn Glu Asn Ser Tyr Gln Met Glu Glu Asp Gly Ser Leu Lys Gln
440 445 450
Ser Ile Leu Ser Ser Glu Leu Leu Asp His Pro Tyr Cys Lys Ser
455 460 465
Pro Leu Glu Ala Pro Leu Val Cys Ser Gly Leu Lys Leu Glu Asn
470 475 480
Gln Val Gly Gly Gly Lys Asn Ser Gln Lys Ala Ser Pro Val Asp
485 490 495
Asp Glu Gln Leu Ser Val Cys Leu Ser Gly Phe Leu Asp Glu Val
500 505 510
Met Lys Lys Tyr Gly Ser Leu Val Pro Leu Ser Glu Lys Glu Val
515 520 525
Leu Gly Arg Leu Lys Asp Val Phe Asn Glu Asp Phe Ser Asn Arg
530 535 540
Lys Pro Phe Ile Asn Arg Glu Ile Thr Asn Tyr Arg Ala Arg His
545 550 555
Gln Lys Cys Asn Phe Arg Ile Phe Tyr Asn Lys His Met Leu Asp
560 565 570
Met Asp Asp Leu Ala Thr Leu Asp Gly Gln Asn Trp Leu Asn Asp
575 580 585
Gln Val Ile Asn Met Tyr Gly Glu Leu Ile Met Asp Ala Val Pro
590 595 600
Asp Lys Val His Phe Phe Asn Ser Phe Phe His Arg Gln Leu Val
605 610 615
Thr Lys Gly Tyr Asn Gly Val Lys Arg Trp Thr Lys Lys Val Asp
620 625 630
Leu Phe Lys Lys Ser Leu Leu Leu Ile Pro Ile His Leu Glu Val
635 640 645
His Trp Ser Leu Ile Thr Val Thr Leu Ser Asn Arg Ile Ile Ser
650 655 660
Phe Tyr Asp Ser Gln Gly Ile His Phe Lys Phe Cys Val Glu Cys

```
                665                     670                     675
Ile Pro Gln Gln Lys Asn Asp Ser Asp Cys Gly Val Phe Val Leu
                680                     685                     690
Gln Tyr Cys Lys Cys Leu Ala Leu Glu Gln Pro Phe Gln Phe Ser
                695                     700                     705
Gln Glu Asp Met Pro Arg Val Arg Lys Arg Ile Tyr Lys Glu Leu
                710                     715                     720
Cys Glu Cys Arg Leu Met Asp
                725

<210> 6
<211> 143
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503356CD1

<400> 6
Met Ser Leu Trp Pro Pro Phe Arg Cys Arg Trp Lys Leu Ala Pro
  1               5                  10                  15
Arg Tyr Ser Arg Arg Ala Ser Pro Gln Gln Pro Gln Gln Asp Phe
                 20                  25                  30
Glu Ala Leu Leu Ala Glu Cys Leu Arg Asn Gly Cys Leu Phe Glu
                 35                  40                  45
Asp Thr Ser Phe Pro Ala Thr Leu Ser Ser Ile Gly Ser Gly Ser
                 50                  55                  60
Leu Leu Gln Lys Leu Pro Pro Arg Leu Gln Trp Lys Arg Pro Pro
                 65                  70                  75
Glu Leu His Ser Asn Pro Gln Phe Tyr Phe Ala Lys Ala Lys Arg
                 80                  85                  90
Leu Asp Leu Cys Gln Gly Ile Val Gly Asp Cys Trp Phe Leu Ala
                 95                 100                 105
Ala Leu Gln Ala Leu Ala Leu His Gln Asp Ile Leu Ser Arg Val
                110                 115                 120
Val Pro Leu Asn Gln Ser Phe Thr Glu Lys Tyr Ala Gly Ile Phe
                125                 130                 135
Arg Phe Trp Ala Leu Trp Phe Leu
                140

<210> 7
<211> 31
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509052CD1

<400> 7
Met Val Asn Pro Thr Val Phe Phe Asp Ile Ala Val Asp Gly Glu
  1               5                  10                  15
Pro Leu Gly Arg Val Ser Phe Glu Val Arg Gly Leu Asp Thr Lys
                 20                  25                  30
Lys

<210> 8
<211> 630
<212> PRT
<213> Homo sapiens
```

<220>
<221> misc_feature
<223> Incyte ID No: 7503366CD1

<400> 8
Met Thr Ala Glu Leu Gln Gln Asp Asp Ala Ala Gly Ala Ala Asp
1 5 10 15
Gly His Gly Ser Ser Cys Gln Met Leu Leu Asn Gln Leu Arg Glu
20 25 30
Ile Thr Gly Ile Gln Asp Pro Ser Phe Leu His Glu Ala Leu Lys
35 40 45
Ala Ser Asn Gly Asp Ile Thr Gln Ala Val Ser Leu Leu Thr Asp
50 55 60
Glu Arg Val Lys Glu Pro Ser Gln Asp Thr Val Ala Thr Glu Pro
65 70 75
Ser Glu Val Glu Gly Ser Ala Ala Asn Lys Glu Val Leu Ala Lys
80 85 90
Val Ile Asp Leu Thr His Asp Asn Lys Asp Asp Leu Gln Ala Ala
95 100 105
Ile Ala Leu Ser Leu Leu Glu Ser Pro Lys Ile Gln Ala Asp Gly
110 115 120
Arg Asp Leu Asn Arg Met His Glu Ala Thr Ser Ala Glu Thr Lys
125 130 135
Arg Ser Lys Arg Lys Arg Cys Glu Val Trp Gly Glu Asn Pro Asn
140 145 150
Pro Asn Asp Trp Arg Arg Val Asp Gly Trp Pro Val Gly Leu Lys
155 160 165
Asn Val Gly Asn Thr Cys Trp Phe Ser Ala Val Ile Gln Ser Leu
170 175 180
Phe Gln Leu Pro Glu Phe Arg Arg Leu Val Leu Ser Tyr Ser Leu
185 190 195
Pro Gln Asn Val Leu Glu Asn Cys Arg Ser His Thr Glu Lys Arg
200 205 210
Asn Ile Met Phe Met Gln Glu Leu Gln Tyr Leu Phe Ala Leu Met
215 220 225
Met Gly Ser Asn Arg Lys Phe Val Asp Pro Ser Ala Ala Leu Asp
230 235 240
Leu Leu Lys Gly Ala Phe Arg Ser Ser Glu Glu Gln Gln Gln Asp
245 250 255
Val Ser Glu Phe Thr His Lys Leu Leu Asp Trp Leu Glu Asp Ala
260 265 270
Phe Gln Leu Ala Val Asn Val Asn Ser Ser Pro Arg Asn Lys Ser
275 280 285
Glu Asn Pro Met Val Gln Leu Phe Tyr Gly Thr Phe Leu Thr Glu
290 295 300
Gly Val Arg Glu Gly Lys Pro Phe Cys Asn Asn Glu Thr Phe Gly
305 310 315
Gln Tyr Pro Leu Gln Val Asn Gly Tyr Arg Asn Leu Asp Glu Cys
320 325 330
Leu Glu Gly Ala Met Val Glu Gly Asp Val Glu Leu Leu Pro Ser
335 340 345
Asp His Ser Val Lys Tyr Gly Gln Glu Arg Trp Phe Thr Lys Leu
350 355 360
Pro Pro Val Leu Thr Phe Glu Leu Ser Arg Phe Glu Phe Asn Gln
365 370 375
Ser Leu Gly Gln Pro Glu Lys Ile His Asn Lys Leu Glu Phe Pro
380 385 390
Gln Ile Ile Tyr Met Asp Arg Tyr Met Tyr Arg Ser Lys Glu Leu
395 400 405
Ile Arg Asn Lys Arg Glu Cys Ile Arg Lys Leu Lys Glu Glu Ile
410 415 420
Lys Ile Leu Gln Gln Lys Leu Glu Arg Tyr Val Lys Tyr Gly Ser
425 430 435

```
Gly Pro Ala Arg Phe Pro Leu Pro Asp Met Leu Lys Tyr Val Ile
                440                 445                 450
Glu Phe Ala Ser Thr Lys Pro Ala Ser Glu Ser Cys Pro Pro Glu
                455                 460                 465
Ser Asp Thr His Met Thr Leu Pro Leu Ser Ser Val His Cys Ser
                470                 475                 480
Val Ser Asp Gln Thr Ser Lys Glu Ser Thr Ser Thr Glu Ser Ser
                485                 490                 495
Ser Gln Asp Val Glu Ser Thr Phe Ser Ser Pro Glu Asp Ser Leu
                500                 505                 510
Pro Lys Ser Lys Pro Leu Thr Ser Ser Arg Ser Ser Met Glu Met
                515                 520                 525
Pro Ser Gln Pro Ala Pro Arg Thr Val Thr Asp Glu Glu Ile Asn
                530                 535                 540
Phe Val Lys Thr Cys Leu Gln Arg Trp Arg Ser Glu Ile Glu Gln
                545                 550                 555
Asp Ile Gln Asp Leu Lys Thr Cys Ile Ala Ser Thr Thr Gln Thr
                560                 565                 570
Ile Glu Gln Met Tyr Cys Asp Pro Leu Leu Arg Gln Phe Leu Phe
                575                 580                 585
Met Lys Asp Lys Gln Met Leu Asp Thr Ile Gly Pro Ile Ser Ile
                590                 595                 600
Ile Asn Pro Asp Arg Ala Gly Ser Ser Thr Met Thr Ser Leu Leu
                605                 610                 615
Leu Asn Leu Pro Gly Lys Lys Leu Lys Glu Ile Pro Met Glu Ala
                620                 625                 630


<210> 9
<211> 354
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505933CD1

<400> 9
Met Val Leu Glu Ser Thr Met Val Cys Val Asp Asn Ser Glu Tyr
 1               5                  10                  15
Met Arg Asn Gly Asp Phe Leu Pro Thr Arg Leu Gln Ala Gln Gln
                20                  25                  30
Asp Ala Val Asn Ile Val Cys His Ser Lys Thr Arg Ser Asn Pro
                35                  40                  45
Glu Asn Asn Val Gly Leu Ile Thr Leu Ala Asn Asp Cys Glu Val
                50                  55                  60
Leu Thr Thr Leu Thr Pro Asp Thr Gly Arg Ile Leu Ser Lys Leu
                65                  70                  75
His Thr Val Gln Pro Lys Gly Lys Ile Thr Phe Cys Thr Gly Ile
                80                  85                  90
Arg Val Ala His Leu Ala Leu Lys His Arg Gln Gly Lys Asn His
                95                  100                 105
Lys Met Arg Ile Ile Ala Phe Val Gly Ser Pro Val Glu Asp Asn
                110                 115                 120
Glu Lys Asp Glu Val Asn Thr Glu Lys Leu Thr Ala Phe Val Asn
                125                 130                 135
Thr Leu Asn Gly Lys Asp Gly Thr Gly Ser His Leu Val Thr Val
                140                 145                 150
Pro Pro Gly Pro Ser Leu Ala Asp Ala Leu Ile Ser Ser Pro Ile
                155                 160                 165
Leu Ala Gly Glu Gly Gly Ala Met Leu Gly Leu Gly Ala Ser Asp
                170                 175                 180
Phe Glu Phe Gly Val Asp Pro Ser Ala Asp Pro Glu Leu Ala Leu
```

```
                185                     190                     195
Ala Leu Arg Val Ser Met Glu Glu Gln Arg Gln Arg Gln Glu Glu
                200                     205                     210
Glu Ala Arg Arg Ala Ala Ala Ala Ser Ala Ala Glu Ala Gly Ile
                215                     220                     225
Ala Thr Thr Gly Thr Glu Asp Ser Asp Asp Ala Leu Leu Lys Met
                230                     235                     240
Thr Ile Ser Gln Gln Glu Phe Gly Arg Thr Gly Leu Pro Asp Leu
                245                     250                     255
Ser Ser Met Thr Glu Glu Glu Gln Ile Ala Tyr Ala Met Gln Met
                260                     265                     270
Ser Leu Gln Gly Ala Glu Phe Gly Gln Ala Glu Ser Ala Asp Ile
                275                     280                     285
Asp Ala Ser Ser Ala Met Asp Thr Ser Glu Pro Ala Lys Glu Glu
                290                     295                     300
Asp Asp Tyr Asp Val Met Gln Asp Pro Glu Phe Leu Gln Ser Val
                305                     310                     315
Leu Glu Asn Leu Pro Gly Val Asp Pro Asn Asn Glu Ala Ile Arg
                320                     325                     330
Asn Ala Met Gly Ser Leu Ala Ser Gln Ala Thr Lys Asp Gly Lys
                335                     340                     345
Lys Asp Lys Lys Glu Glu Asp Lys Lys
                350
```

<210> 10
<211> 132
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507064CD1

<400> 10

```
Met Gln Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met
  1               5                  10                  15
Ser Arg Asn Asp Gln Gly Ser Met Leu Thr Leu Arg Ala Ile Asp
                 20                  25                  30
Leu Ser Tyr Tyr Thr Gly Ser Leu His Trp Ile Pro Met Thr Ala
                 35                  40                  45
Arg Ile Leu Ala Val His Cys Gly Gln Phe Asp Ile Asp Cys Gly
                 50                  55                  60
Arg Leu Ser Ser Ile Pro Thr Ala Val Phe Glu Ile His Gly Lys
                 65                  70                  75
Lys Tyr Pro Leu Pro Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly
                 80                  85                  90
Phe Cys Thr Ser Gly Phe Gln Gly Asp Tyr Ser Ser Gln Gln Trp
                 95                 100                 105
Ile Leu Gly Asn Val Phe Ile Trp Glu Tyr Tyr Ser Val Phe Asp
                110                 115                 120
Arg Thr Asn Asn Arg Val Gly Leu Ala Lys Ala Val
                125                 130
```

<210> 11
<211> 316
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1439986CD1

<400> 11

Met Gly Gln Asp Tyr Tyr Ser Val Leu Gly Ile Thr Arg Asn Ser
1 5 10 15
Glu Asp Ala Gln Ile Lys Gln Ala Tyr Arg Arg Leu Ala Leu Lys
20 25 30
His His Pro Leu Lys Ser Asn Glu Pro Ser Ser Ala Glu Ile Phe
35 40 45
Arg Gln Ile Ala Glu Ala Tyr Asp Val Leu Ser Asp Pro Met Lys
50 55 60
Arg Gly Ile Tyr Asp Lys Phe Gly Glu Glu Gly Leu Lys Gly Gly
65 70 75
Ile Pro Leu Glu Phe Gly Ser Gln Thr Pro Trp Thr Thr Gly Tyr
80 85 90
Val Phe His Gly Lys Pro Glu Lys Val Phe His Glu Phe Phe Gly
95 100 105
Gly Asn Asn Pro Phe Ser Glu Phe Phe Asp Ala Glu Gly Ser Glu
110 115 120
Val Asp Leu Asn Phe Gly Gly Leu Gln Gly Arg Gly Val Lys Lys
125 130 135
Gln Asp Pro Gln Val Glu Arg Asp Leu Tyr Leu Ser Leu Glu Asp
140 145 150
Leu Phe Phe Gly Cys Thr Lys Lys Ile Lys Ile Ser Arg Arg Val
155 160 165
Leu Asn Glu Asp Gly Tyr Ser Ser Thr Ile Lys Asp Lys Ile Leu
170 175 180
Thr Ile Asp Val Lys Pro Gly Trp Arg Gln Gly Thr Arg Ile Thr
185 190 195
Phe Glu Lys Glu Gly Asp Gln Gly Pro Asn Ile Ile Pro Ala Asp
200 205 210
Ile Ile Phe Ile Val Lys Glu Lys Leu His Pro Arg Phe Arg Arg
215 220 225
Glu Asn Asp Asn Leu Phe Phe Val Asn Pro Ile Pro Leu Gly Lys
230 235 240
Ala Leu Thr Cys Cys Thr Val Glu Val Arg Thr Leu Asp Asp Arg
245 250 255
Leu Leu Asn Ile Pro Ile Asn Asp Ile Ile His Pro Lys Tyr Phe
260 265 270
Lys Lys Val Pro Gly Glu Gly Met Pro Leu Pro Glu Asp Pro Thr
275 280 285
Lys Lys Gly Asp Leu Phe Ile Phe Phe Asp Ile Gln Phe Pro Thr
290 295 300
Arg Leu Thr Pro Gln Lys Lys Gln Met Leu Arg Gln Ala Leu Leu
305 310 315
Thr

<210> 12
<211> 531
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2008979CD1

<400> 12
Met Lys Cys His Tyr Glu Ala Leu Gly Val Arg Arg Asp Ala Ser
1 5 10 15
Glu Glu Glu Leu Lys Lys Ala Tyr Arg Lys Lèu Ala Leu Lys Trp
20 25 30
His Pro Asp Lys Asn Leu Asp Asn Ala Ala Glu Ala Ala Glu Gln
35 40 45
Phe Lys Leu Ile Gln Ala Ala Tyr Asp Val Leu Ser Asp Pro Gln
50 55 60

Glu Arg Ala Trp Tyr Asp Asn His Arg Glu Ala Leu Leu Lys Gly
65 70 75
Gly Phe Asp Gly Glu Tyr Gln Asp Asp Ser Leu Asp Leu Leu Arg
80 85 90
Tyr Phe Thr Val Thr Cys Tyr Ser Gly Tyr Gly Asp Asp Glu Lys
95 100 105
Gly Phe Tyr Thr Val Tyr Arg Asn Val Phe Glu Met Ile Ala Lys
110 115 120
Glu Glu Leu Glu Ser Val Leu Glu Glu Glu Val Asp Asp Phe Pro
125 130 135
Thr Phe Gly Asp Ser Gln Ser Asp Tyr Asp Thr Val Val His Pro
140 145 150
Phe Tyr Ala Tyr Trp Gln Ser Phe Cys Thr Gln Lys Asn Phe Ala
155 160 165
Trp Lys Glu Glu Tyr Asp Thr Arg Gln Ala Ser Asn Arg Trp Glu
170 175 180
Lys Arg Ala Met Glu Lys Glu Asn Lys Lys Ile Arg Asp Lys Ala
185 190 195
Arg Lys Glu Lys Asn Glu Leu Val Arg Gln Leu Val Ala Phe Ile
200 205 210
Arg Lys Arg Asp Lys Arg Val Gln Ala His Arg Lys Leu Val Glu
215 220 225
Glu Gln Asn Ala Glu Lys Ala Arg Lys Ala Glu Glu Met Arg Arg
230 235 240
Gln Gln Lys Leu Lys Gln Ala Lys Leu Val Glu Gln Tyr Arg Glu
245 250 255
Gln Ser Trp Met Thr Met Ala Asn Leu Glu Lys Glu Leu Gln Glu
260 265 270
Met Glu Ala Arg Tyr Glu Lys Glu Phe Gly Asp Gly Ser Asp Glu
275 280 285
Asn Glu Met Glu Glu His Glu Leu Lys Asp Glu Glu Asp Gly Lys
290 295 300
Asp Ser Asp Glu Ala Glu Asp Ala Glu Leu Tyr Asp Asp Leu Tyr
305 310 315
Cys Pro Ala Cys Asp Lys Ser Phe Lys Thr Glu Lys Ala Met Lys
320 325 330
Asn His Glu Lys Ser Lys Lys His Arg Glu Met Val Ala Leu Leu
335 340 345
Lys Gln Gln Leu Glu Glu Glu Glu Glu Asn Phe Ser Arg Pro Gln
350 355 360
Ile Asp Glu Asn Pro Leu Asp Asp Asn Ser Glu Glu Glu Met Glu
365 370 375
Asp Ala Pro Lys Gln Lys Leu Ser Lys Lys Gln Lys Lys Lys Lys
380 385 390
Gln Lys Pro Ala Gln Asn Tyr Asp Asp Asn Phe Asn Val Asn Gly
395 400 405
Pro Gly Glu Gly Val Lys Val Asp Pro Glu Asp Thr Asn Leu Asn
410 415 420
Gln Asp Ser Ala Lys Glu Leu Glu Asp Ser Pro Gln Glu Asn Val
425 430 435
Ser Val Thr Glu Ile Ile Lys Pro Cys Asp Asp Pro Lys Ser Glu
440 445 450
Ala Lys Ser Val Pro Lys Pro Lys Gly Lys Lys Thr Lys Asp Met
455 460 465
Lys Lys Pro Val Arg Val Pro Ala Glu Pro Gln Thr Met Ser Val
470 475 480
Leu Ile Ser Cys Thr Thr Cys His Ser Glu Phe Pro Ser Arg Asn
485 490 495
Lys Leu Phe Asp His Leu Lys Ala Thr Gly His Ala Arg Ala Pro
500 505 510
Ser Ser Ser Ser Leu Asn Ser Ala Thr Ser Ser Gln Ser Lys Lys
515 520 525
Glu Lys Arg Lys Asn Arg

530

<210> 13
<211> 108
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 90073157CD1

<400> 13
Met Ser Arg Pro Gln Leu Arg Arg Trp Arg Leu Val Ser Ser Pro
1 5 10 15
Pro Ser Gly Val Pro Gly Leu Ala Leu Leu Ala Leu Leu Ala Leu
20 25 30
Leu Ala Leu Arg Leu Ala Ala Gly Thr Asp Cys Pro Cys Pro Glu
35 40 45
Pro Glu Leu Cys Arg Pro Ile Arg His His Pro Asp Phe Glu Val
50 55 60
Phe Val Phe Asp Val Gly Gln Lys Thr Trp Lys Ser Tyr Asp Trp
65 70 75
Ser Gln Ile Thr Thr Val Ala Thr Phe Gly Lys Tyr Asp Ser Glu
80 85 90
Leu Met Cys Tyr Ala His Ser Lys Gly Ala Arg Val Val Leu Lys
95 100 105
Gly Asn Leu

<210> 14
<211> 345
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506782CD1

<400> 14
Met Ala Phe Leu Gly Leu Phe Ser Leu Leu Val Leu Gln Ser Met
1 5 10 15
Ala Thr Gly Ala Thr Phe Pro Glu Glu Ala Ile Ala Asp Leu Ser
20 25 30
Val Asn Met Tyr Asn Arg Leu Arg Ala Thr Gly Glu Asp Glu Asn
35 40 45
Ile Leu Phe Ser Pro Leu Ser Ile Ala Leu Ala Met Gly Met Met
50 55 60
Glu Leu Gly Ala Gln Gly Ser Thr Gln Lys Glu Ile Arg His Ser
65 70 75
Met Gly Tyr Asp Ser Leu Lys Asn Gly Glu Glu Phe Ser Phe Leu
80 85 90
Lys Glu Phe Ser Asn Met Val Thr Ala Lys Glu Ser Gln Tyr Val
95 100 105
Met Lys Ile Ala Asn Ser Leu Phe Val Gln Asn Gly Phe His Val
110 115 120
Asn Glu Glu Phe Leu Gln Met Met Lys Lys Tyr Phe Asn Ala Ala
125 130 135
Val Asn His Val Asp Phe Ser Gln Asn Val Ala Val Ala Asn Tyr
140 145 150
Ile Asn Lys Trp Val Glu Asn Asn Thr Asn Arg Glu Phe Ser Asp
155 160 165
Gly Ser Asn Glu Ala Gly Gly Ile Tyr Gln Val Leu Glu Ile Pro
170 175 180

Tyr Glu Gly Asp Glu Ile Ser Met Met Leu Val Leu Ser Arg Gln
185 190 195
Glu Val Pro Leu Ala Thr Leu Glu Pro Leu Val Lys Ala Gln Leu
200 205 210
Val Glu Glu Trp Ala Asn Ser Val Lys Lys Gln Lys Val Glu Val
215 220 225
Tyr Leu Pro Arg Phe Thr Val Glu Gln Glu Ile Asp Leu Lys Asp
230 235 240
Val Leu Lys Ala Leu Gly Ile Thr Glu Ile Phe Ile Lys Asp Ala
245 250 255
Asn Leu Thr Gly Leu Ser Asp Asn Lys Glu Ile Phe Leu Ser Lys
260 265 270
Ala Ile His Lys Ser Phe Leu Glu Val Asn Glu Glu Gly Ser Glu
275 280 285
Ala Ala Ala Val Ser Gly Met Ile Ala Ile Ser Arg Met Ala Val
290 295 300
Leu Tyr Pro Gln Val Ile Val Asp His Pro Phe Phe Phe Leu Ile
305 310 315
Arg Asn Arg Arg Thr Gly Thr Ile Leu Phe Met Gly Arg Val Met
320 325 330
His Pro Glu Thr Met Asn Thr Ser Gly His Asp Phe Glu Glu Leu
335 340 345

<210> 15
<211> 493
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506941CD1

<400> 15
Met Pro Asp Gln Leu Glu Ser Leu Pro Leu Phe Ser Lys Lys Asn
1 5 10 15
Leu Ile Asp Ile Leu Ala Ile Gly Gly Val Gln Lys Leu Lys Gln
20 25 30
Leu Pro Val Val Val Lys Phe Leu Glu Phe Tyr Met Lys Lys Met
35 40 45
Met Asn Leu Arg Trp Lys Leu Phe Met Leu His Pro Leu Cys Trp
50 55 60
Lys Gln Gly Arg Ala Asp Ser Phe Arg Tyr Pro Lys Thr Gly Thr
65 70 75
Ala Asn Pro Lys Val Thr Phe Lys Met Ser Glu Ile Met Ile Asp
80 85 90
Ala Glu Gly Arg Ile His Asp Ile Phe His Val Phe Pro Gln Ser
95 100 105
His Glu Glu Glu Ile Glu Phe Ile Phe Ala Ser Glu Cys Lys Thr
110 115 120
Gly Phe Arg His Leu Tyr Lys Ile Thr Ser Ile Leu Lys Glu Ser
125 130 135
Lys Tyr Lys Arg Ser Ser Gly Gly Leu Pro Ala Pro Ser Asp Phe
140 145 150
Lys Cys Pro Ile Lys Glu Glu Ile Ala Ile Thr Ser Gly Glu Trp
155 160 165
Glu Val Leu Gly Arg His Gly Ser Asn Ile Gln Val Asp Glu Val
170 175 180
Arg Arg Leu Val Tyr Phe Glu Gly Thr Lys Asp Ser Pro Leu Glu
185 190 195
His His Leu Tyr Val Val Ser Tyr Val Asn Pro Gly Glu Val Thr
200 205 210
Arg Leu Thr Asp Arg Gly Tyr Ser His Ser Cys Cys Ile Ser Gln

```
                    215                     220                     225
His Cys Asp Phe Phe Ile Ser Lys Tyr Ser Asn Gln Lys Asn Pro
                    230                     235                     240
His Cys Val Ser Leu Tyr Lys Leu Ser Ser Pro Glu Asp Asp Pro
                    245                     250                     255
Thr Cys Lys Thr Lys Glu Phe Trp Ala Thr Ile Leu Asp Ser Ala
                    260                     265                     270
Gly Pro Leu Pro Asp Tyr Thr Pro Pro Glu Ile Phe Ser Phe Glu
                    275                     280                     285
Ser Thr Thr Gly Phe Thr Leu Tyr Gly Met Leu Tyr Lys Pro His
                    290                     295                     300
Asp Leu Gln Pro Gly Lys Lys Tyr Pro Thr Val Leu Phe Ile Tyr
                    305                     310                     315
Gly Gly Pro Gln Val Gln Leu Val Asn Asn Arg Phe Lys Gly Val
                    320                     325                     330
Lys Tyr Phe Arg Leu Asn Thr Leu Ala Ser Leu Gly Tyr Val Val
                    335                     340                     345
Val Val Ile Asp Asn Arg Gly Ser Cys His Arg Gly Leu Lys Phe
                    350                     355                     360
Glu Gly Ala Phe Lys Tyr Lys Met Val Ala Ile Ala Gly Ala Pro
                    365                     370                     375
Val Thr Leu Trp Ile Phe Tyr Asp Thr Gly Tyr Thr Glu Arg Tyr
                    380                     385                     390
Met Gly His Pro Asp Gln Asn Glu Gln Gly Tyr Tyr Leu Gly Ser
                    395                     400                     405
Val Ala Met Gln Ala Glu Lys Phe Pro Ser Glu Pro Asn Arg Leu
                    410                     415                     420
Leu Leu Leu His Gly Phe Leu Asp Glu Asn Val His Phe Ala His
                    425                     430                     435
Thr Ser Ile Leu Leu Ser Phe Leu Val Arg Ala Gly Lys Pro Tyr
                    440                     445                     450
Asp Leu Gln Ile Tyr Pro Gln Glu Arg His Ser Ile Arg Val Pro
                    455                     460                     465
Glu Ser Gly Glu His Tyr Glu Leu His Leu Leu His Tyr Leu Gln
                    470                     475                     480
Glu Asn Leu Gly Ser Arg Ile Ala Ala Leu Lys Val Ile
                    485                     490
```

<210> 16
<211> 204
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507072CD1

<400> 16

```
Met Gln Gly Thr Pro Gly Gly Gly Thr Arg Pro Gly Pro Ser Pro
 1               5                  10                  15
Val Asp Arg Arg Thr Leu Leu Val Phe Ser Phe Ile Leu Ala Ala
                20                  25                  30
Ala Leu Gly Gln Met Asn Phe Thr Gly Asp Gln Val Leu Arg Val
                35                  40                  45
Leu Ala Lys Asp Glu Lys Gln Leu Ser Leu Leu Gly Asp Leu Glu
                50                  55                  60
Gly Leu Lys Pro Gln Lys Val Asp Phe Trp Arg Gly Pro Ala Arg
                65                  70                  75
Pro Ser Leu Pro Val Asp Met Arg Val Pro Phe Ser Glu Leu Lys
                80                  85                  90
Asp Ile Lys Ala Tyr Leu Glu Ser His Gly Leu Ala Tyr Ser Ile
                95                 100                 105
Met Ile Lys Asp Ile Gln Val Leu Leu Asp Glu Glu Arg Arg Ala
```

110 115 120
Met Ala Lys Ser Arg Arg Leu Glu Arg Ser Thr Asn Ser Phe Ser
125 130 135
Tyr Ser Ser Tyr His Thr Leu Glu Glu Ile Tyr Ser Trp Ile Asp
140 145 150
Asn Phe Val Met Glu His Ser Asp Ile Val Ser Lys Ile Gln Ile
155 160 165
Gly Asn Ser Phe Glu Asn Gln Ser Ile Leu Val Leu Lys Glu Met
170 175 180
Val Leu Thr Ala Thr Pro Ala Gln Lys Leu Ile Thr Gly Pro Pro
185 190 195
Leu Ser Arg Ser Arg Arg Trp Leu Pro
200

<210> 17
<211> 224
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507083CD1

<400> 17
Met Ala Asp Glu Lys Pro Ser Asn Gly Val Leu Val His Met Val
1 5 10 15
Lys Leu Leu Ile Lys Thr Phe Leu Asp Gly Ile Phe Asp Asp Leu
20 25 30
Met Glu Asn Asn Val Leu Asn Thr Asp Glu Ile His Leu Ile Gly
35 40 45
Lys Cys Leu Lys Phe Val Val Ser Asn Ala Glu Asn Leu Val Asp
50 55 60
Asp Ile Thr Glu Thr Ala Gln Ile Ala Gly Lys Ile Phe Arg Glu
65 70 75
His Leu Trp Asn Ser Lys Lys Gln Leu Ser Ser Asp Val Ser Ser
80 85 90
Asp Gly Glu Arg Glu Ala Asn Met Pro Gly Leu Asn Ile Arg Asn
95 100 105
Lys Glu Phe Asn Tyr Leu His Asn Arg Asn Gly Ser Glu Leu Asp
110 115 120
Leu Leu Gly Met Arg Asp Leu Leu Glu Asn Leu Gly Tyr Ser Val
125 130 135
Val Ile Lys Glu Asn Leu Thr Ala Gln Met Val Leu Gly Leu Phe
140 145 150
Gly Ser Pro Leu Thr Val Glu Lys Pro Val Gln Ile Leu Met Val
155 160 165
Gly Ser Cys Lys Val Thr Ser Val Met Met Leu Leu Gln Arg Leu
170 175 180
Met Trp Lys Arg Thr Ser Leu Leu Ser Asn Leu Pro His His Ile
185 190 195
Met Phe Leu Gly Asp Met Lys Gln Met Ala Pro Ser Ser Phe Pro
200 205 210
Lys Leu Ser Thr Thr Ser Glu Ser Ile Leu Gly Val Ile Ile
215 220

<210> 18
<211> 277
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509097CD1

<400> 18
Met Pro Glu Lys Pro Gly Pro Arg Trp Ile Phe Ser Ser Thr Ser
1   5   10   15
Ala Lys Pro Gly Gly Leu Phe Pro Ser Glu Thr Met Glu Arg Asp
                20   25   30
Ser His Gly Asn Ala Ser Pro Ala Arg Thr Pro Ser Ala Gly Ala
                35   40   45
Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr Pro Pro Gly Arg Ala
                50   55   60
Ser Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr
                65   70   75
Pro Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr
                80   85   90
Pro Pro Gly Arg Ala Ser Pro Gly Arg Ala Ser Pro Ala Gln Ala
                95   100   105
Ser Pro Ala Gln Ala Ser Pro Ala Arg Ala Ser Pro Ala Leu Ala
                110   115   120
Ser Leu Ser Arg Ser Ser Ser Gly Arg Ser Ser Ser Ala Arg Ser
                125   130   135
Ala Ser Val Thr Thr Ser Pro Thr Arg Val Tyr Leu Val Arg Ala
                140   145   150
Thr Pro Val Gly Ala Val Pro Ile Arg Ser Ser Pro Ala Arg Ser
                155   160   165
Ala Pro Ala Thr Arg Ala Thr Arg Glu Ser Pro Gly Thr Ser Leu
                170   175   180
Pro Lys Phe Thr Trp Arg Glu Gly Gln Lys Gln Leu Pro Leu Ile
                185   190   195
Gly Cys Val Leu Leu Leu Ile Ala Leu Val Val Ser Leu Ile Ile
                200   205   210
Leu Phe Gln Phe Trp Gln Gly His Thr Gly Ile Arg Tyr Lys Glu
                215   220   225
Gln Arg Glu Ser Cys Pro Lys His Ala Val Arg Cys Asp Gly Val
                230   235   240
Val Asp Cys Lys Leu Lys Ser Asp Glu Leu Gly Cys Val Arg Phe
                245   250   255
Asp Trp Asp Lys Ser Leu Leu Lys Ile Tyr Ser Gly Ser Ser His
                260   265   270
Gln Trp Leu Pro Thr Ala Asp
                275

<210> 19
<211> 449
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509118CD1

<400> 19
Met Pro Glu Lys Pro Gly Pro Arg Trp Ile Phe Ser Ser Thr Ser
1   5   10   15
Ala Lys Pro Gly Gly Leu Phe Pro Ser Glu Thr Met Glu Arg Asp
                20   25   30
Ser His Gly Asn Ala Ser Pro Ala Arg Thr Pro Ser Ala Gly Ala
                35   40   45
Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr Pro Pro Gly Arg Ala
                50   55   60
Ser Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr
                65   70   75
Pro Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr
                80   85   90
Pro Pro Gly Arg Ala Ser Pro Gly Arg Ala Ser Pro Ala Gln Ala

```
                95                      100                     105
Ser Pro Ala Gln Ala Ser Pro Ala Arg Ala Ser Pro Ala Leu Ala
                110                     115                     120
Ser Leu Ser Arg Ser Ser Ser Gly Arg Ser Ser Ser Ala Arg Ser
                125                     130                     135
Ala Ser Val Thr Thr Ser Pro Thr Arg Val Tyr Leu Val Arg Ala
                140                     145                     150
Thr Pro Val Gly Ala Val Pro Ile Arg Ser Ser Pro Ala Arg Ser
                155                     160                     165
Ala Pro Ala Thr Arg Ala Thr Arg Glu Ser Pro Gly Thr Ser Leu
                170                     175                     180
Pro Lys Phe Thr Trp Arg Glu Gly Gln Lys Gln Leu Pro Leu Ile
                185                     190                     195
Gly Cys Val Leu Leu Leu Ile Ala Leu Val Val Ser Leu Ile Ile
                200                     205                     210
Leu Phe Gln Phe Trp Gln Gly His Thr Gly Ile Arg Tyr Lys Glu
                215                     220                     225
Gln Arg Glu Ser Cys Pro Lys His Ala Val Arg Cys Asp Gly Val
                230                     235                     240
Val Asp Cys Lys Leu Lys Ser Asp Glu Leu Gly Cys Val Arg Phe
                245                     250                     255
Asp Trp Asp Lys Ser Leu Leu Lys Ile Tyr Ser Gly Ser Ser His
                260                     265                     270
Gln Trp Leu Pro Ile Cys Ser Ser Asn Trp Asn Asp Ser Tyr Ser
                275                     280                     285
Glu Lys Thr Cys Gln Gln Leu Gly Phe Glu Ser Ala His Arg Thr
                290                     295                     300
Thr Glu Val Ala His Arg Asp Phe Ala Asn Ser Phe Ser Ile Leu
                305                     310                     315
Arg Tyr Asn Ser Thr Ile Gln Glu Ser Leu His Arg Ser Glu Cys
                320                     325                     330
Pro Ser Gln Arg Tyr Ile Ser Leu Gln Cys Ser His Cys Gly Leu
                335                     340                     345
Arg Ala Met Thr Gly Arg Ile Val Gly Gly Ala Leu Ala Ser Asp
                350                     355                     360
Ser Lys Trp Pro Trp Gln Val Ser Leu His Phe Gly Thr Thr His
                365                     370                     375
Ile Cys Gly Gly Thr Leu Ile Asp Ala Gln Trp Val Leu Thr Ala
                380                     385                     390
Ala His Cys Phe Phe Val Thr Arg Glu Lys Val Leu Glu Gly Trp
                395                     400                     405
Lys Val Tyr Ala Gly Thr Ser Asn Ser Tyr Pro Gly Pro Lys Ala
                410                     415                     420
Ser Ala Gly Gln Lys Ser Lys Thr Leu Lys Asp Pro Tyr Met Glu
                425                     430                     435
His Phe Cys Phe Ile Ile Arg Glu Thr Glu Ala Gln Gly Leu
                440                     445

<210> 20
<211> 451
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509312CD1

<400> 20
Met Glu Arg Asp Ser His Gly Asn Ala Ser Pro Ala Arg Thr Pro
  1             5                   10                      15
Ser Ala Gly Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr Pro
                20                  25                      30
Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser
```

35 40 45
Pro Ala Arg Ala Ser Pro Ala Leu Ala Ser Leu Ser Arg Ser Ser
50 55 60
Ser Gly Arg Ser Ser Ser Ala Arg Ser Ala Ser Val Thr Thr Ser
65 70 75
Pro Thr Arg Val Tyr Leu Val Arg Ala Thr Pro Val Gly Ala Val
80 85 90
Pro Ile Arg Ser Ser Pro Ala Arg Ser Ala Pro Ala Thr Arg Ala
95 100 105
Thr Arg Glu Ser Pro Gly Thr Ser Leu Pro Lys Phe Thr Trp Arg
110 115 120
Glu Gly Gln Lys Gln Leu Pro Leu Ile Gly Cys Val Leu Leu Leu
125 130 135
Ile Ala Leu Val Val Ser Leu Ile Ile Leu Phe Gln Phe Trp Gln
140 145 150
Gly His Thr Gly Ile Arg Tyr Lys Glu Gln Arg Glu Ser Cys Pro
155 160 165
Lys His Ala Val Arg Cys Asp Gly Val Val Asp Cys Lys Leu Lys
170 175 180
Ser Asp Glu Leu Gly Cys Val Arg Phe Asp Trp Asp Lys Ser Leu
185 190 195
Leu Lys Ile Tyr Ser Gly Ser Ser His Gln Trp Leu Pro Ile Cys
200 205 210
Ser Ser Asn Trp Asn Asp Ser Tyr Ser Glu Lys Thr Cys Gln Gln
215 220 225
Leu Gly Phe Glu Ser Ala His Arg Thr Thr Glu Val Ala His Arg
230 235 240
Asp Phe Ala Asn Ser Phe Ser Ile Leu Arg Tyr Asn Ser Thr Ile
245 250 255
Gln Glu Ser Leu His Arg Ser Glu Cys Pro Ser Gln Arg Tyr Ile
260 265 270
Ser Leu Gln Cys Ser His Cys Gly Leu Arg Ala Met Thr Gly Arg
275 280 285
Ile Val Gly Gly Ala Leu Ala Ser Asp Ser Lys Trp Pro Trp Gln
290 295 300
Val Ser Leu His Phe Gly Thr Thr His Ile Cys Gly Gly Thr Leu
305 310 315
Ile Asp Ala Gln Trp Val Leu Thr Ala Ala His Cys Phe Phe Val
320 325 330
Thr Arg Glu Lys Val Leu Glu Gly Trp Lys Val Tyr Ala Gly Thr
335 340 345
Ser Asn Leu His Gln Leu Pro Glu Ala Ala Ser Ile Ala Glu Ile
350 355 360
Ile Ile Asn Ser Asn Tyr Thr Asp Glu Glu Asp Asp Tyr Asp Ile
365 370 375
Ala Leu Met Arg Leu Ser Lys Pro Leu Thr Leu Ser Gly Glu Gly
380 385 390
Ile Cys Thr Pro Arg Ser Pro Ala Pro Gln Pro Gln His Pro Leu
395 400 405
Gln Pro Ser His Leu Ser Ala Ser Val Asn Ser Tyr Pro Gly Pro
410 415 420
Lys Ala Ser Ala Gly Gln Lys Ser Lys Thr Leu Lys Asp Pro Tyr
425 430 435
Met Glu His Phe Cys Phe Ile Ile Arg Glu Thr Glu Ala Gln Gly
440 445 450
Leu

<210> 21
<211> 485
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 90126902CD1

<400> 21
Met Gly Lys Leu Arg Pro Gly Arg Val Glu Trp Leu Ala Ser Gly
1 5 10 15
His Thr Glu Arg Pro His Leu Phe Gln Asn Leu Leu Leu Phe Leu
20 25 30
Trp Ala Leu Leu Asn Cys Gly Leu Gly Val Ser Ala Gln Gly Pro
35 40 45
Gly Glu Trp Thr Pro Trp Val Ser Trp Thr Arg Cys Ser Ser Ser
50 55 60
Cys Gly Arg Gly Val Ser Val Arg Ser Arg Arg Cys Leu Arg Leu
65 70 75
Pro Gly Glu Glu Pro Cys Trp Gly Asp Ser His Glu Tyr Arg Leu
80 85 90
Cys Gln Leu Pro Asp Cys Pro Pro Gly Ala Val Pro Phe Arg Asp
95 100 105
Leu Gln Cys Ala Leu Tyr Asn Gly Arg Pro Val Leu Gly Thr Gln
110 115 120
Lys Thr Tyr Gln Trp Val Pro Phe His Gly Ala Pro Asn Gln Cys
125 130 135
Asp Leu Asn Cys Leu Ala Glu Gly His Ala Phe Tyr His Ser Phe
140 145 150
Gly Arg Val Leu Asp Gly Thr Ala Cys Ser Pro Gly Ala Gln Gly
155 160 165
Val Cys Val Ala Gly Arg Cys Leu Ser Ala Gly Cys Asp Gly Leu
170 175 180
Leu Gly Ser Gly Ala Leu Glu Asp Arg Cys Gly Arg Cys Gly Gly
185 190 195
Ala Asn Asp Ser Cys Leu Phe Val Gln Arg Val Phe Arg Asp Ala
200 205 210
Gly Ala Phe Ala Gly Tyr Trp Asn Val Thr Leu Ile Pro Glu Gly
215 220 225
Ala Arg His Ile Arg Val Glu His Arg Ser Arg Asn His Leu Gly
230 235 240
Ile Leu Gly Ser Leu Met Gly Gly Asp Gly Arg Tyr Val Leu Asn
245 250 255
Gly His Trp Val Val Ser Pro Pro Gly Thr Tyr Glu Ala Ala Gly
260 265 270
Thr His Val Val Tyr Thr Arg Asp Thr Gly Pro Gln Glu Thr Leu
275 280 285
Gln Ala Ala Gly Pro Thr Ser His Asp Leu Leu Leu Gln Val Leu
290 295 300
Leu Gln Glu Pro Asn Pro Gly Ile Glu Phe Glu Phe Trp Leu Pro
305 310 315
Arg Glu Arg Tyr Ser Pro Phe Gln Ala Arg Val Gln Ala Leu Gly
320 325 330
Trp Pro Leu Arg Gln Pro Gln Pro Arg Gly Val Glu Pro Gln Pro
335 340 345
Pro Ala Ala Pro Ala Val Thr Pro Ala Gln Thr Pro Thr Leu Ala
350 355 360
Pro Asp Pro Cys Pro Pro Cys Pro Asp Thr Arg Gly Arg Ala His
365 370 375
Arg Leu Leu His Tyr Cys Gly Ser Asp Phe Val Phe Gln Ala Arg
380 385 390
Val Leu Gly His His His Gln Ala Gln Glu Thr Arg Tyr Glu Val
395 400 405
Arg Ile Gln Leu Val Tyr Lys Asn Arg Ser Pro Leu Arg Ala Arg
410 415 420
Glu Tyr Val Trp Ala Pro Gly His Cys Pro Cys Pro Met Leu Ala
425 430 435

```
Pro His Arg Asp Tyr Leu Met Ala Val Gln Arg Leu Val Ser Pro
                440                 445                 450
Asp Gly Thr Gln Asp Gln Leu Leu Leu Pro His Ala Gly Tyr Ala
                455                 460                 465
Arg Pro Trp Ser Pro Ala Glu Asp Ser Arg Ile Arg Leu Thr Ala
                470                 475                 480
Arg Arg Cys Pro Gly
                485
```

<210> 22
<211> 258
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509352CD1

<400> 22

```
Met Ala Trp Ser Pro Pro Ala Thr Leu Phe Leu Phe Leu Leu Leu
  1               5                  10                  15
Leu Gly Gln Pro Pro Pro Ser Arg Pro Gln Ser Leu Gly Thr Thr
                 20                  25                  30
Lys Leu Arg Leu Val Gly Pro Glu Ser Lys Pro Glu Glu Gly Arg
                 35                  40                  45
Leu Glu Val Leu His Gln Gly Gln Trp Gly Thr Val Cys Asp Asp
                 50                  55                  60
Asn Phe Ala Ile Gln Glu Ala Thr Val Ala Cys Arg Gln Leu Gly
                 65                  70                  75
Phe Glu Ala Ala Leu Thr Trp Ala His Ser Ala Lys Tyr Gly Gln
                 80                  85                  90
Gly Glu Gly Pro Ile Trp Leu Asp Asn Val Arg Cys Val Gly Thr
                 95                 100                 105
Glu Ser Ser Leu Asp Gln Cys His Arg His Tyr His Ser Ile Glu
                110                 115                 120
Val Phe Thr His Tyr Asp Leu Leu Thr Leu Asn Gly Ser Lys Val
                125                 130                 135
Ala Glu Gly His Lys Ala Ser Phe Cys Leu Glu Asp Thr Asn Cys
                140                 145                 150
Pro Thr Gly Leu Gln Arg Arg Tyr Ala Cys Ala Asn Phe Gly Glu
                155                 160                 165
Gln Gly Val Thr Val Gly Cys Trp Asp Thr Tyr Arg His Asp Ile
                170                 175                 180
Asp Cys Gln Trp Val Asp Ile Thr Asp Val Gly Pro Gly Asn Tyr
                185                 190                 195
Ile Phe Gln Val Ile Val Asn Pro His Tyr Glu Val Ala Glu Ser
                200                 205                 210
Asp Phe Ser Asn Asn Met Leu Gln Cys Arg Cys Lys Tyr Asp Gly
                215                 220                 225
His Arg Val Trp Leu His Asn Cys His Thr Gly Asn Ser Tyr Pro
                230                 235                 240
Ala Asn Ala Glu Leu Ser Leu Glu Gln Glu Gln Arg Leu Arg Asn
                245                 250                 255
Asn Leu Ile
```

<210> 23
<211> 252
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7509341CD1

<400> 23
Met Gly Leu Arg Ala Gly Pro Ile Leu Leu Leu Leu Leu Trp Leu
1   5   10   15
Leu Pro Gly Ala His Trp Asp Val Leu Pro Ser Glu Cys Gly His
                20   25   30
Ser Lys Glu Ala Gly Arg Ile Val Gly Gly Gln Asp Thr Gln Glu
                35   40   45
Gly Arg Trp Pro Trp Gln Val Gly Leu Trp Leu Thr Ser Val Gly
                50   55   60
His Val Cys Gly Gly Ser Leu Ile His Pro Arg Trp Val Leu Thr
                65   70   75
Ala Ala His Cys Phe Leu Arg Ser Glu Asp Pro Gly Leu Tyr His
                80   85   90
Val Lys Val Gly Gly Leu Thr Pro Ser Leu Ser Glu Pro His Ser
                95   100   105
Ala Leu Val Ala Val Arg Arg Leu Leu Val His Ser Ser Tyr His
                110   115   120
Gly Thr Thr Thr Ser Gly Asp Ile Ala Leu Met Glu Leu Asp Ser
                125   130   135
Pro Leu Gln Ala Ser Gln Phe Ser Pro Ile Cys Leu Pro Gly Pro
                140   145   150
Gln Thr Pro Leu Ala Ile Gly Thr Val Cys Trp Gly Asp Ser Gly
                155   160   165
Gly Pro Leu Val Cys Pro Ile Asn Asp Thr Trp Ile Gln Ala Gly
                170   175   180
Ile Val Ser Trp Gly Phe Gly Cys Ala Arg Pro Phe Arg Pro Gly
                185   190   195
Val Tyr Thr Gln Val Leu Ser Tyr Thr Asp Trp Ile Gln Arg Thr
                200   205   210
Leu Ala Glu Ser His Ser Gly Met Ser Gly Ala Arg Pro Gly Ala
                215   220   225
Pro Gly Ser His Ser Gly Thr Ser Arg Ser His Pro Val Leu Leu
                230   235   240
Leu Glu Leu Leu Thr Val Cys Leu Leu Gly Ser Leu
                245   250

<210> 24
<211> 102
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509367CD1

<400> 24
Met Gly Leu Arg Ala Gly Pro Ile Leu Leu Leu Leu Leu Trp Leu
1   5   10   15
Leu Pro Gly Ala His Trp Asp Val Leu Pro Ser Glu Cys Gly His
                20   25   30
Ser Lys Glu Ala Gly Arg Ile Val Gly Gly Gln Asp Thr Gln Glu
                35   40   45
Gly Arg Trp Pro Trp Gln Val Gly Leu Trp Leu Thr Ser Val Gly
                50   55   60
His Val Cys Gly Gly Ser Leu Ile His Pro Arg Trp Val Leu Thr
                65   70   75
Ala Ala His Cys Phe Leu Arg Val Thr Pro Gly Gly Arg Trp Ser
                80   85   90
Ala Pro Ser Met Ile Arg Gly Ser Arg Pro Ala Leu
                95   100

<210> 25
<211> 62
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500455CD1

<400> 25
Met Pro Ala Val Ala Ser Val Pro Lys Glu Leu Tyr Leu Ser Ser
1   5   10   15
Ser Leu Lys Asp Leu Asn Lys Lys Thr Glu Val Lys Pro Glu Lys
                20   25   30
Ile Ser Thr Lys Arg Ile Ile Ser Ile Gln Tyr Leu Asp Leu Glu
                35   40   45
Thr Ser Lys Lys Leu Ser Lys Lys Leu Lys Asp Ser Leu Lys Ala
                50   55   60
Leu Asn

<210> 26
<211> 181
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510401CD1

<400> 26
Met Ala Glu Leu Thr Ala Leu Glu Ser Leu Ile Glu Met Gly Phe
1   5   10   15
Pro Arg Gly Arg Ala Glu Lys Ala Leu Ala Leu Thr Gly Asn Gln
                20   25   30
Gly Ile Glu Ala Ala Met Asp Trp Leu Met Glu His Glu Asp Asp
                35   40   45
Pro Asp Val Asp Glu Pro Leu Glu Thr Pro Leu Gly His Ile Leu
                50   55   60
Gly Arg Glu Pro Thr Ser Ser Glu Gln Gly Gly Leu Glu Gly Ser
                65   70   75
Gly Ser Ala Ala Gly Glu Gly Lys Pro Ala Leu Ser Glu Glu Glu
                80   85   90
Arg Gln Glu Gln Thr Lys Arg Met Leu Glu Leu Val Ala Gln Lys
                95   100   105
Gln Arg Glu Ser Glu Glu Arg Glu Glu Arg Glu Ala Leu Glu Arg
                110   115   120
Glu Arg Gln Arg Arg Arg Gln Gly Gln Glu Leu Ser Ala Ala Arg
                125   130   135
Gln Arg Leu Gln Glu Asp Glu Met Arg Arg Ala Ala Glu Glu Arg
                140   145   150
Arg Arg Glu Lys Ala Glu Glu Leu Ala Ala Arg Gln Arg Val Arg
                155   160   165
Glu Lys Ile Glu Arg Asp Lys Ala Glu Arg Ala Lys Lys Val Gly
                170   175   180
Asp

<210> 27
<211> 104
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504702CD1

<400> 27
```
Met Ser Pro Pro Pro Leu Leu Gln Pro Leu Leu Leu Leu Leu Pro
1               5                   10                  15
Leu Leu Asn Val Glu Pro Ser Gly Ala Thr Leu Ile Arg Ile Pro
                20                  25                  30
Leu His Arg Val Gln Pro Gly Arg Arg Ile Leu Asn Leu Leu Arg
                35                  40                  45
Gly Trp Arg Glu Pro Ala Glu Leu Pro Lys Leu Gly Ala Pro Ser
                50                  55                  60
Pro Gly Asp Lys Pro Ile Phe Val Pro Leu Ser Asn Tyr Arg Asp
                65                  70                  75
Gly Tyr Thr Thr Asp Leu Ile Pro Lys Pro Leu Ala Pro Ser Arg
                80                  85                  90
Pro Met Gly Pro Ser Leu Pro Phe Asn Met Glu Leu Gly Gly
                95                  100
```

<210> 28
<211> 444
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509113CD1

<400> 28
```
Met Gly Arg Pro Leu His Leu Val Leu Leu Ser Ala Ser Leu Ala
1               5                   10                  15
Gly Leu Leu Leu Leu Gly Glu Ser Leu Phe Ile Arg Arg Glu Gln
                20                  25                  30
Ala Asn Asn Ile Leu Ala Arg Val Thr Arg Ala Asn Ser Phe Leu
                35                  40                  45
Glu Glu Met Lys Lys Gly His Leu Glu Arg Glu Cys Met Glu Glu
                50                  55                  60
Thr Cys Ser Tyr Glu Glu Ala Arg Glu Val Phe Glu Asp Ser Asp
                65                  70                  75
Lys Thr Asn Glu Phe Trp Asn Lys Tyr Lys Asp Gly Asp Gln Cys
                80                  85                  90
Glu Thr Ser Pro Cys Gln Asn Gln Gly Lys Cys Lys Asp Gly Leu
                95                  100                 105
Gly Glu Tyr Thr Cys Thr Cys Leu Glu Gly Phe Glu Gly Lys Asn
                110                 115                 120
Cys Glu Leu Trp Pro Tyr Pro Cys Gly Lys Gln Thr Leu Glu Arg
                125                 130                 135
Arg Lys Arg Ser Val Ala Gln Ala Thr Ser Ser Ser Gly Glu Ala
                140                 145                 150
Pro Asp Ser Ile Thr Trp Lys Pro Tyr Asp Ala Ala Asp Leu Asp
                155                 160                 165
Pro Thr Glu Asn Pro Phe Asp Leu Leu Asp Phe Asn Gln Thr Gln
                170                 175                 180
Pro Glu Arg Gly Asp Asn Asn Leu Thr Arg Ile Val Gly Gly Gln
                185                 190                 195
Glu Cys Lys Asp Gly Glu Cys Pro Trp Gln Ala Leu Leu Ile Asn
                200                 205                 210
Glu Glu Asn Glu Gly Phe Cys Gly Gly Thr Ile Leu Ser Glu Phe
                215                 220                 225
Tyr Ile Leu Thr Ala Ala His Cys Leu Tyr Gln Ala Lys Arg Phe
                230                 235                 240
Lys Val Arg Val Gly Asp Arg Asn Thr Glu Gln Glu Glu Gly Gly
```

```
                245                     250                     255
Glu Ala Val His Glu Val Glu Val Val Ile Lys His Asn Arg Phe
                260                     265                     270
Thr Lys Glu Thr Tyr Asp Phe Asp Ile Ala Val Leu Arg Leu Lys
                275                     280                     285
Thr Pro Ile Thr Phe Arg Met Asn Val Ala Pro Ala Cys Leu Pro
                290                     295                     300
Glu Arg Asp Trp Ala Glu Ser Thr Leu Met Thr Gln Lys Thr Gly
                305                     310                     315
Ile Val Ser Gly Phe Gly Arg Thr His Glu Lys Gly Arg Gln Ser
                320                     325                     330
Thr Arg Leu Lys Met Leu Glu Val Pro Tyr Val Asp Arg Asn Ser
                335                     340                     345
Cys Lys Leu Ser Ser Ser Phe Ile Ile Thr Gln Asn Met Phe Cys
                350                     355                     360
Ala Gly Tyr Asp Thr Lys Gln Glu Asp Ala Cys Gln Gly Asp Ser
                365                     370                     375
Gly Gly Pro His Val Thr Arg Phe Lys Asp Thr Tyr Phe Val Thr
                380                     385                     390
Gly Ile Val Ser Trp Gly Glu Gly Cys Ala Arg Lys Gly Lys Tyr
                395                     400                     405
Gly Ile Tyr Thr Lys Val Thr Ala Phe Leu Lys Trp Ile Asp Arg
                410                     415                     420
Ser Met Lys Thr Arg Gly Leu Pro Lys Ala Lys Ser His Ala Pro
                425                     430                     435
Glu Val Ile Thr Ser Ser Pro Leu Lys
                440
```

<210> 29
<211> 377
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509140CD1

<400> 29

```
Met Gly Arg Pro Leu His Leu Val Leu Leu Ser Ala Ser Leu Ala
  1               5                      10                      15
Gly Leu Leu Leu Leu Gly Glu Ser Leu Phe Ile Arg Arg Glu Gln
                 20                      25                      30
Ala Asn Asn Ile Leu Ala Arg Val Thr Arg Ala Asn Ser Phe Leu
                 35                      40                      45
Glu Glu Met Lys Lys Gly His Leu Glu Arg Glu Cys Met Glu Glu
                 50                      55                      60
Thr Cys Ser Tyr Glu Glu Ala Arg Glu Val Phe Glu Asp Ser Asp
                 65                      70                      75
Lys Thr Asn Glu Phe Trp Asn Lys Tyr Lys Asp Gly Asp Gln Cys
                 80                      85                      90
Glu Thr Ser Pro Cys Gln Asn Gln Gly Lys Cys Lys Asp Gly Leu
                 95                     100                     105
Gly Glu Tyr Thr Cys Thr Cys Leu Glu Gly Phe Glu Gly Lys Asn
                110                     115                     120
Cys Glu Leu Phe Thr Arg Lys Leu Cys Ser Leu Asp Asn Gly Asp
                125                     130                     135
Cys Asp Gln Phe Cys His Glu Glu Gln Asn Ser Val Val Cys Ser
                140                     145                     150
Cys Ala Arg Gly Tyr Thr Leu Ala Asp Asn Gly Lys Ala Cys Ile
                155                     160                     165
Pro Thr Gly Pro Tyr Pro Cys Gly Lys Gln Thr Leu Glu Arg Arg
                170                     175                     180
Lys Arg Ser Val Ala Gln Ala Thr Ser Ser Ser Gly Glu Ala Pro
```

```
                185                 190                 195
Asp Ser Ile Thr Trp Lys Pro Tyr Asp Ala Ala Asp Leu Asp Pro
                200                 205                 210
Thr Glu Asn Pro Phe Asp Leu Leu Asp Phe Asn Gln Thr Gln Pro
                215                 220                 225
Glu Arg Gly Asp Asn Asn Leu Thr Arg Ile Val Gly Gly Gln Glu
                230                 235                 240
Cys Lys Asp Gly Glu Cys Pro Trp Gln Ala Leu Leu Ile Asn Glu
                245                 250                 255
Glu Asn Glu Lys Gly Arg Gln Ser Thr Arg Leu Lys Met Leu Glu
                260                 265                 270
Val Pro Tyr Val Asp Arg Asn Ser Cys Lys Leu Ser Ser Ser Phe
                275                 280                 285
Ile Ile Thr Gln Asn Met Phe Cys Ala Gly Tyr Asp Thr Lys Gln
                290                 295                 300
Glu Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Arg
                305                 310                 315
Phe Lys Asp Thr Tyr Phe Val Thr Gly Ile Val Ser Trp Gly Glu
                320                 325                 330
Gly Cys Ala Arg Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Thr
                335                 340                 345
Ala Phe Leu Lys Trp Ile Asp Arg Ser Met Lys Thr Arg Gly Leu
                350                 355                 360
Pro Lys Ala Lys Ser His Ala Pro Glu Val Ile Thr Ser Ser Pro
                365                 370                 375
Leu Lys
```

<210> 30
<211> 442
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509223CD1

<400> 30

```
Met Gly Arg Pro Leu His Leu Val Leu Leu Ser Ala Ser Leu Ala
 1               5                  10                  15
Gly Leu Leu Leu Leu Gly Glu Ser Leu Phe Ile Arg Arg Glu Gln
                20                  25                  30
Ala Asn Asn Ile Leu Ala Arg Val Thr Arg Ala Asn Ser Phe Leu
                35                  40                  45
Glu Glu Met Lys Lys Gly His Leu Glu Arg Glu Cys Met Glu Glu
                50                  55                  60
Thr Cys Ser Tyr Glu Glu Ala Arg Glu Val Phe Glu Asp Ser Asp
                65                  70                  75
Lys Thr Asn Glu Phe Trp Asn Lys Tyr Lys Asp Gly Asp Gln Cys
                80                  85                  90
Glu Thr Ser Pro Cys Gln Asn Gln Gly Lys Cys Lys Asp Gly Leu
                95                 100                 105
Gly Glu Tyr Thr Cys Thr Cys Leu Glu Gly Phe Glu Gly Lys Asn
               110                 115                 120
Cys Glu Leu Phe Thr Arg Lys Leu Cys Ser Leu Asp Asn Gly Asp
               125                 130                 135
Cys Asp Gln Phe Cys His Glu Glu Gln Asn Ser Val Val Cys Ser
               140                 145                 150
Cys Ala Arg Gly Tyr Thr Leu Ala Asp Asn Gly Lys Ala Cys Ile
               155                 160                 165
Pro Thr Gly Pro Tyr Pro Cys Gly Lys Gln Thr Leu Glu Arg Arg
               170                 175                 180
Lys Arg Ser Val Ala Gln Ala Thr Ser Ser Ser Gly Glu Ala Pro
```

```
                    185                     190                 195
Asp Ser Ile Thr Trp Lys Pro Tyr Asp Ala Ala Asp Leu Asp Pro
                    200                     205                 210
Thr Glu Asn Pro Phe Asp Leu Leu Asp Phe Asn Gln Thr Gln Pro
                    215                     220                 225
Glu Arg Gly Asp Asn Asn Leu Thr Arg Ile Val Gly Gly Gln Glu
                    230                     235                 240
Cys Lys Asp Gly Glu Cys Pro Trp Gln Ala Leu Leu Ile Asn Glu
                    245                     250                 255
Glu Asn Glu Val Glu Val Val Ile Lys His Asn Arg Phe Thr Lys
                    260                     265                 270
Glu Thr Tyr Asp Phe Asp Ile Ala Val Leu Arg Leu Lys Thr Pro
                    275                     280                 285
Ile Thr Phe Arg Met Asn Val Ala Pro Ala Cys Leu Pro Glu Arg
                    290                     295                 300
Asp Trp Ala Glu Ser Thr Leu Met Thr Gln Lys Thr Gly Ile Val
                    305                     310                 315
Ser Gly Phe Gly Arg Thr His Glu Lys Gly Arg Gln Ser Thr Arg
                    320                     325                 330
Leu Lys Met Leu Glu Val Pro Tyr Val Asp Arg Asn Ser Cys Lys
                    335                     340                 345
Leu Ser Ser Ser Phe Ile Ile Thr Gln Asn Met Phe Cys Ala Gly
                    350                     355                 360
Tyr Asp Thr Lys Gln Glu Asp Ala Cys Gln Gly Asp Ser Gly Gly
                    365                     370                 375
Pro His Val Thr Arg Phe Lys Asp Thr Tyr Phe Val Thr Gly Ile
                    380                     385                 390
Val Ser Trp Gly Glu Gly Cys Ala Arg Lys Gly Lys Tyr Gly Ile
                    395                     400                 405
Tyr Thr Lys Val Thr Ala Phe Leu Lys Trp Ile Asp Arg Ser Met
                    410                     415                 420
Lys Thr Arg Gly Leu Pro Lys Ala Lys Ser His Ala Pro Glu Val
                    425                     430                 435
Ile Thr Ser Ser Pro Leu Lys
                    440
```

<210> 31
<211> 375
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509272CD1

<400> 31

```
Met Glu Arg Asp Ser His Gly Asn Ala Ser Pro Ala Arg Thr Pro
  1               5                  10                  15
Ser Ala Gly Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr Pro
                 20                  25                  30
Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser
                 35                  40                  45
Pro Ala Gly Thr Pro Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser
                 50                  55                  60
Pro Ala Gly Thr Pro Pro Gly Arg Ala Ser Pro Gly Arg Ala Ser
                 65                  70                  75
Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser Pro Ala Arg Ala Ser
                 80                  85                  90
Pro Ala Leu Ala Ser Leu Ser Arg Ser Ala Pro Ala Thr Arg Ala
                 95                 100                 105
Thr Arg Glu Ser Pro Gly Thr Ser Leu Pro Lys Phe Thr Trp Arg
                110                 115                 120
Glu Gly Gln Lys Gln Leu Pro Leu Ile Gly Cys Val Leu Leu Leu
```

```
                    125                     130                     135
Ile Ala Leu Val Val Ser Leu Ile Ile Leu Phe Gln Phe Trp Gln
                140                 145                 150
Gly His Thr Gly Ile Arg Tyr Lys Glu Gln Arg Glu Ser Cys Pro
                155                 160                 165
Lys His Ala Val Arg Cys Asp Gly Val Val Asp Cys Lys Leu Lys
                170                 175                 180
Ser Asp Glu Leu Gly Cys Val Arg Phe Asp Trp Asp Lys Ser Leu
                185                 190                 195
Leu Lys Ile Tyr Ser Gly Ser Ser His Gln Trp Leu Pro Ile Cys
                200                 205                 210
Ser Ser Asn Trp Asn Asp Ser Tyr Ser Glu Lys Thr Cys Gln Gln
                215                 220                 225
Leu Gly Phe Glu Ser Ala His Arg Thr Thr Glu Val Ala His Arg
                230                 235                 240
Asp Phe Ala Asn Ser Phe Ser Ile Leu Arg Tyr Asn Ser Thr Ile
                245                 250                 255
Gln Glu Ser Leu His Arg Ser Glu Cys Pro Ser Gln Arg Tyr Ile
                260                 265                 270
Ser Leu Gln Cys Ser His Cys Gly Leu Arg Ala Met Thr Gly Arg
                275                 280                 285
Ile Val Gly Gly Ala Leu Ala Ser Asp Ser Lys Trp Pro Trp Gln
                290                 295                 300
Val Ser Leu His Phe Gly Thr Thr His Ile Cys Gly Gly Thr Leu
                305                 310                 315
Ile Asp Ala Gln Trp Val Leu Thr Ala Ala His Cys Phe Phe Val
                320                 325                 330
Thr Arg Glu Lys Val Leu Glu Gly Trp Lys Val Tyr Ala Gly Thr
                335                 340                 345
Ser Asn Leu His Gln Leu Pro Glu Ala Ala Ser Ile Ala Glu Ile
                350                 355                 360
Ile Ile Asn Ser Asn Tyr Thr Asp Glu Glu Glu Val Lys Asp Ser
                365                 370                 375
```

```
<210> 32
<211> 204
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509327CD1

<400> 32
Met Glu Arg Asp Ser His Gly Asn Ala Ser Pro Ala Arg Thr Pro
  1             5                   10                  15
Ser Ala Gly Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr Pro
                20                  25                  30
Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser
                35                  40                  45
Pro Ala Gly Thr Pro Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser
                50                  55                  60
Pro Ala Gly Thr Pro Pro Gly Arg Ala Ser Pro Gly Arg Ala Ser
                65                  70                  75
Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser Pro Ala Arg Ala Ser
                80                  85                  90
Pro Ala Leu Ala Ser Leu Ser Arg Ser Ser Ser Gly Arg Ser Ser
                95                  100                 105
Ser Ala Arg Ser Ala Ser Val Thr Thr Ser Pro Thr Arg Val Tyr
                110                 115                 120
Leu Val Arg Ala Thr Pro Val Gly Ala Val Pro Ile Arg Ser Ser
                125                 130                 135
```

Pro Ala Arg Ser Ala Pro Ala Thr Arg Ala Thr Arg Glu Ser Pro
140 145 150
Gly Thr Ser Leu Pro Lys Phe Thr Trp Arg Glu Gly Gln Lys Gln
155 160 165
Leu Pro Leu Ile Gly Cys Val Leu Leu Leu Ile Ala Leu Val Val
170 175 180
Ser Leu Ile Ile Leu Leu Leu Thr Gly Gln Pro Arg Leu Pro Thr
185 190 195
Gly Ile Leu Pro Thr Ala Ser Gln Ser
200

<210> 33
<211> 86
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504677CD1

<400> 33
Met Pro Lys Thr Met His Phe Leu Phe Arg Phe Ile Val Phe Phe
1 5 10 15
Tyr Leu Trp Gly Leu Phe Thr Ala Gln Arg Gln Lys Lys Glu Glu
20 25 30
Ser Thr Glu Glu Val Lys Ile Glu Val Leu His Arg Pro Glu Asn
35 40 45
Cys Ser Lys Thr Ser Lys Lys Gly Asp Leu Leu Asn Ala His Tyr
50 55 60
Asp Gly Tyr Leu Ala Lys Asp Gly Ser Lys Phe Tyr Cys Ser Arg
65 70 75
Arg Gln Asp Ser Thr Gly Cys Tyr Ile Asp Phe
80 85

<210> 34
<211> 882
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504534CD1

<400> 34
Met Ala Glu Gly Gly Gly Cys Arg Glu Arg Pro Asp Ala Glu Thr
1 5 10 15
Gln Lys Ser Glu Leu Gly Pro Leu Met Arg Thr Thr Leu Gln Arg
20 25 30
Gly Ala Gln Trp Tyr Leu Ile Asp Ser Arg Trp Phe Lys Gln Trp
35 40 45
Lys Lys Tyr Val Gly Phe Asp Ser Trp Gly Met Tyr Asn Val Gly
50 55 60
Glu His Asn Leu Phe Pro Gly Pro Ile Asp Asn Ser Gly Leu Phe
65 70 75
Ser Asp Pro Glu Ser Gln Thr Leu Lys Glu His Leu Ile Asp Glu
80 85 90
Leu Asp Tyr Val Leu Val Pro Thr Glu Ala Trp Asn Lys Leu Leu
95 100 105
Asn Trp Tyr Gly Cys Val Glu Gly Gln Gln Pro Ile Val Arg Lys
110 115 120
Val Val Glu His Gly Leu Phe Val Lys His Cys Lys Val Glu Val
125 130 135
Tyr Leu Leu Glu Leu Lys Leu Cys Glu Asn Ser Asp Pro Thr Asn

140 145 150
Val Leu Ser Cys His Phe Ser Lys Ala Asp Thr Ile Ala Thr Ile
155 160 165
Glu Lys Glu Met Arg Lys Leu Phe Asn Ile Pro Ala Glu Arg Glu
170 175 180
Thr Arg Leu Trp Asn Lys Tyr Met Ser Asn Thr Tyr Glu Gln Leu
185 190 195
Ser Lys Leu Asp Asn Thr Val Gln Asp Ala Gly Leu Tyr Gln Gly
200 205 210
Gln Val Leu Val Ile Glu Pro Gln Asn Glu Asp Gly Thr Trp Pro
215 220 225
Arg Gln Thr Leu Gln Ser Lys Ser Ser Thr Ala Pro Ser Arg Asn
230 235 240
Phe Thr Thr Ser Pro Lys Ser Ser Ala Ser Pro Tyr Ser Ser Val
245 250 255
Ser Ala Ser Leu Ile Ala Asn Gly Asp Ser Thr Ser Thr Cys Gly
260 265 270
Met His Ser Ser Gly Val Ser Arg Gly Gly Ser Gly Phe Ser Ala
275 280 285
Ser Tyr Asn Cys Gln Glu Pro Pro Ser Ser His Ile Gln Pro Gly
290 295 300
Leu Cys Gly Leu Gly Asn Leu Gly Asn Thr Cys Phe Met Asn Ser
305 310 315
Ala Leu Gln Cys Leu Ser Asn Thr Ala Pro Leu Thr Asp Tyr Phe
320 325 330
Leu Lys Asp Glu Tyr Glu Ala Glu Ile Asn Arg Asp Asn Pro Leu
335 340 345
Gly Met Lys Gly Glu Ile Ala Glu Ala Tyr Ala Glu Leu Ile Lys
350 355 360
Gln Met Trp Ser Gly Arg Asp Ala His Val Ala Pro Arg Met Phe
365 370 375
Lys Thr Gln Val Gly Arg Phe Ala Pro Gln Phe Ser Gly Tyr Gln
380 385 390
Gln Gln Asp Ser Gln Glu Leu Leu Ala Phe Leu Leu Asp Gly Leu
395 400 405
His Glu Asp Leu Asn Arg Val Lys Lys Lys Pro Tyr Leu Glu Leu
410 415 420
Lys Asp Ala Asn Gly Arg Pro Asp Ala Val Val Ala Lys Glu Ala
425 430 435
Trp Glu Asn His Arg Leu Arg Asn Asp Ser Val Ile Val Asp Thr
440 445 450
Phe His Gly Leu Phe Lys Ser Thr Leu Val Cys Pro Glu Cys Ala
455 460 465
Lys Val Ser Val Thr Phe Asp Pro Phe Cys Tyr Leu Thr Leu Pro
470 475 480
Leu Pro Leu Lys Lys Asp Arg Val Met Glu Val Phe Leu Val Pro
485 490 495
Ala Asp Pro His Cys Arg Pro Thr Gln Tyr Arg Val Thr Val Pro
500 505 510
Leu Met Gly Ala Val Ser Asp Leu Cys Glu Ala Leu Ser Arg Leu
515 520 525
Ser Gly Ile Ala Ala Glu Asn Met Val Val Ala Asp Val Tyr Asn
530 535 540
His Arg Phe His Lys Ile Phe Gln Met Asp Glu Gly Leu Asn His
545 550 555
Ile Met Pro Arg Asp Asp Ile Phe Val Tyr Glu Val Cys Ser Thr
560 565 570
Ser Val Asp Gly Ser Glu Cys Val Thr Leu Pro Val Tyr Phe Arg
575 580 585
Glu Arg Lys Ser Arg Pro Ser Ser Thr Ser Ser Ala Ser Ala Leu
590 595 600
Tyr Gly Gln Pro Leu Leu Leu Ser Val Pro Lys His Lys Leu Thr
605 610 615

Leu Glu Ser Leu Tyr Gln Ala Val Cys Asp Arg Ile Ser Arg Tyr
620 625 630
Val Lys Gln Pro Leu Pro Asp Glu Phe Gly Ser Ser Pro Leu Glu
635 640 645
Pro Gly Ala Cys Asn Gly Ser Arg Asn Ser Cys Glu Gly Glu Asp
650 655 660
Glu Glu Glu Met Glu His Gln Glu Glu Gly Lys Glu Gln Leu Ser
665 670 675
Glu Thr Glu Gly Ser Gly Glu Asp Glu Pro Gly Asn Asp Pro Ser
680 685 690
Glu Thr Thr Gln Lys Lys Ile Lys Gly Gln Pro Cys Pro Lys Arg
695 700 705
Leu Phe Thr Phe Ser Leu Val Asn Ser Tyr Gly Thr Ala Asp Ile
710 715 720
Asn Ser Leu Ala Ala Asp Gly Lys Leu Leu Lys Leu Asn Ser Arg
725 730 735
Ser Thr Leu Ala Met Asp Trp Asp Ser Glu Thr Arg Arg Leu Tyr
740 745 750
Tyr Asp Glu Gln Glu Ser Glu Ala Tyr Glu Lys His Val Ser Met
755 760 765
Leu Gln Pro Gln Lys Lys Lys Lys Thr Thr Val Ala Leu Arg Asp
770 775 780
Cys Ile Glu Leu Phe Thr Thr Met Glu Thr Leu Gly Glu His Asp
785 790 795
Pro Trp Tyr Cys Pro Asn Cys Lys Lys His Gln Gln Ala Thr Lys
800 805 810
Lys Phe Asp Leu Trp Ser Leu Pro Lys Ile Leu Val Val His Leu
815 820 825
Lys Arg Phe Ser Tyr Asn Arg Tyr Trp Arg Asp Lys Leu Asp Thr
830 835 840
Val Val Glu Phe Pro Ile Arg Gly Leu Asn Met Ser Glu Phe Val
845 850 855
Cys Asn Leu Ser Ala Arg Pro Tyr Val Tyr Asp Leu Ile Ala Val
860 865 870
Ser Asn His Tyr Gly Ala Met Gly Val Gly His Tyr
875 880

<210> 35
<211> 4374
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507771CD1

<400> 35
Met Lys Val Asp Arg Thr Lys Leu Lys Lys Thr Pro Thr Glu Ala
1 5 10 15
Pro Ala Asp Cys Arg Ala Leu Ile Asp Lys Leu Lys Val Cys Asn
20 25 30
Asp Glu Gln Leu Leu Leu Glu Leu Gln Gln Ile Lys Thr Trp Asn
35 40 45
Ile Gly Lys Cys Glu Leu Tyr His Trp Val Asp Leu Leu Asp Arg
50 55 60
Phe Asp Gly Ile Leu Ala Asp Ala Gly Gln Thr Val Glu Asn Met
65 70 75
Ser Trp Met Leu Val Cys Asp Arg Pro Glu Arg Glu Gln Leu Lys
80 85 90
Met Leu Leu Leu Ala Val Leu Asn Phe Thr Ala Leu Leu Ile Glu
95 100 105
Tyr Ser Phe Ser Arg His Leu Tyr Ser Ser Ile Glu His Leu Thr
110 115 120

Thr Leu Leu Ala Ser Ser Asp Met Gln Val Val Leu Ala Val Leu
125 130 135
Asn Leu Leu Tyr Val Phe Ser Lys Arg Ser Asn Tyr Ile Thr Arg
140 145 150
Leu Gly Ser Asp Lys Arg Thr Pro Leu Leu Thr Arg Leu Gln His
155 160 165
Leu Ala Glu Ser Trp Gly Gly Lys Glu Asn Gly Phe Gly Leu Ala
170 175 180
Glu Cys Cys Arg Asp Leu His Met Met Lys Tyr Pro Pro Ser Ala
185 190 195
Thr Thr Leu His Phe Glu Phe Tyr Ala Asp Pro Gly Ala Glu Val
200 205 210
Lys Ile Glu Lys Arg Thr Thr Ser Asn Thr Leu His Tyr Ile His
215 220 225
Ile Glu Gln Leu Asp Lys Ile Ser Glu Ser Pro Ser Glu Ile Met
230 235 240
Glu Ser Leu Thr Lys Met Tyr Ser Thr Pro Lys Asp Lys Gln Met
245 250 255
Leu Leu Phe Thr His Ile Arg Leu Ala His Gly Phe Ser Asn His
260 265 270
Arg Lys Arg Leu Gln Ala Val Gln Ala Arg Leu His Ala Ile Ser
275 280 285
Ile Leu Val Tyr Ser Asn Ala Leu Gln Glu Ser Ala Asn Ser Ile
290 295 300
Leu Tyr Asn Gly Leu Ile Glu Glu Leu Val Asp Val Leu Gln Ile
305 310 315
Thr Asp Lys Gln Leu Met Glu Ile Lys Ala Ala Ser Leu Arg Thr
320 325 330
Leu Thr Ser Ile Val His Leu Glu Arg Thr Pro Lys Leu Ser Ser
335 340 345
Ile Ile Asp Cys Thr Gly Thr Ala Ser Tyr His Gly Phe Leu Pro
350 355 360
Val Leu Val Arg Asn Cys Ile Gln Ala Met Ile Asp Pro Ser Met
365 370 375
Asp Pro Tyr Pro His Gln Phe Ala Thr Ala Leu Phe Ser Phe Leu
380 385 390
Tyr His Leu Ala Ser Tyr Asp Ala Gly Gly Glu Ala Leu Val Ser
395 400 405
Cys Gly Met Met Glu Ala Leu Leu Lys Val Ile Lys Phe Leu Gly
410 415 420
Asp Glu Gln Asp Gln Ile Thr Phe Val Thr Arg Ala Val Arg Val
425 430 435
Val Asp Leu Ile Thr Asn Leu Asp Met Ala Ala Phe Gln Ser His
440 445 450
Ser Gly Leu Ser Ile Phe Ile Tyr Arg Leu Glu His Glu Val Asp
455 460 465
Leu Cys Arg Lys Glu Cys Pro Phe Val Ile Lys Pro Lys Ile Gln
470 475 480
Arg Pro Asn Thr Thr Gln Glu Gly Glu Glu Met Glu Thr Asp Met
485 490 495
Asp Gly Val Gln Cys Ile Pro Gln Arg Ala Ala Leu Leu Lys Ser
500 505 510
Met Leu Asn Phe Leu Lys Lys Ala Ile Gln Asp Pro Ala Phe Ser
515 520 525
Asp Gly Ile Arg His Val Met Asp Gly Ser Leu Pro Thr Ser Leu
530 535 540
Lys His Ile Ile Ser Asn Ala Glu Tyr Tyr Gly Pro Ser Leu Phe
545 550 555
Leu Leu Ala Thr Glu Val Val Thr Val Phe Val Phe Gln Glu Pro
560 565 570
Ser Leu Leu Ser Ser Leu Gln Asp Asn Gly Leu Thr Asp Val Met
575 580 585
Leu His Ala Leu Leu Ile Lys Asp Val Pro Ala Thr Arg Glu Val

590 595 600
Leu Gly Ser Leu Pro Asn Val Phe Ser Ala Leu Cys Leu Asn Ala
605 610 615
Arg Gly Leu Gln Ser Phe Val Gln Cys Gln Pro Phe Glu Arg Leu
620 625 630
Phe Lys Val Leu Leu Ser Pro Asp Tyr Leu Pro Ala Met Arg Arg
635 640 645
Arg Arg Ser Ser Asp Pro Leu Gly Asp Thr Ala Ser Asn Leu Gly
650 655 660
Ser Ala Val Asp Glu Leu Met Arg His Gln Pro Thr Leu Lys Thr
665 670 675
Asp Ala Thr Thr Ala Ile Ile Lys Leu Leu Glu Glu Ile Cys Asn
680 685 690
Leu Gly Arg Asp Pro Lys Tyr Ile Cys Gln Lys Pro Ser Ile Gln
695 700 705
Lys Ala Asp Gly Thr Ala Thr Ala Pro Pro Pro Arg Ser Asn His
710 715 720
Ala Ala Glu Glu Ala Ser Ser Glu Asp Glu Glu Glu Glu Glu Val
725 730 735
Gln Ala Met Gln Ser Phe Asn Ser Thr Gln Gln Asn Glu Thr Glu
740 745 750
Pro Asn Gln Gln Val Val Gly Thr Glu Glu Arg Ile Pro Ile Pro
755 760 765
Leu Met Asp Tyr Ile Leu Asn Val Met Lys Phe Val Glu Ser Ile
770 775 780
Leu Ser Asn Asn Thr Thr Asp Asp His Cys Gln Glu Phe Val Asn
785 790 795
Gln Lys Gly Leu Leu Pro Leu Val Thr Ile Leu Gly Leu Pro Asn
800 805 810
Leu Pro Ile Asp Phe Pro Thr Ser Ala Ala Cys Gln Ala Val Ala
815 820 825
Gly Val Cys Lys Ser Ile Leu Thr Leu Ser His Glu Pro Lys Val
830 835 840
Leu Gln Glu Gly Leu Leu Gln Leu Asp Ser Ile Leu Ser Ser Leu
845 850 855
Glu Pro Leu His Arg Pro Ile Glu Ser Pro Gly Gly Ser Val Leu
860 865 870
Leu Arg Glu Leu Ala Cys Ala Gly Asn Val Ala Asp Ala Thr Leu
875 880 885
Ser Ala Gln Ala Thr Pro Leu Leu His Ala Leu Thr Ala Ala His
890 895 900
Ala Tyr Ile Met Met Phe Val His Thr Cys Arg Val Gly Gln Ser
905 910 915
Glu Ile Arg Ser Ile Ser Val Asn Gln Trp Gly Ser Gln Leu Gly
920 925 930
Leu Ser Val Leu Ser Lys Leu Ser Gln Leu Tyr Cys Ser Leu Val
935 940 945
Trp Glu Ser Thr Val Leu Leu Ser Leu Cys Thr Pro Asn Ser Leu
950 955 960
Pro Ser Gly Cys Glu Phe Gly Gln Ala Asp Met Gln Lys Leu Val
965 970 975
Pro Lys Asp Glu Lys Ala Gly Thr Thr Gln Gly Gly Lys Arg Ser
980 985 990
Asp Gly Glu Gln Asp Gly Ala Ala Gly Ser Met Asp Ala Ser Thr
995 1000 1005
Gln Gly Leu Leu Glu Gly Ile Gly Leu Asp Gly Asp Thr Leu Ala
1010 1015 1020
Pro Met Glu Thr Asp Glu Pro Thr Ala Ser Asp Ser Lys Gly Lys
1025 1030 1035
Ser Lys Ile Thr Pro Ala Met Ala Ala Arg Ile Lys Gln Ile Lys
1040 1045 1050
Pro Leu Leu Ser Ala Ser Ser Arg Leu Gly Arg Ala Leu Ala Glu
1055 1060 1065

```
Leu Phe Gly Leu Leu Val Lys Leu Cys Val Gly Ser Pro Val Arg
                1070                1075                1080
Gln Arg Arg Ser His His Ala Ala Ser Thr Thr Thr Ala Pro Thr
                1085                1090                1095
Pro Ala Ala Arg Ser Thr Ala Ser Ala Leu Thr Lys Leu Leu Thr
                1100                1105                1110
Lys Gly Leu Ser Trp Gln Pro Pro Pro Tyr Thr Pro Thr Pro Arg
                1115                1120                1125
Phe Arg Leu Thr Phe Phe Ile Cys Ser Val Gly Phe Thr Ser Pro
                1130                1135                1140
Met Leu Phe Asp Glu Arg Lys Tyr Pro Tyr His Leu Met Leu Gln
                1145                1150                1155
Lys Phe Leu Cys Ser Gly Gly His Asn Ala Leu Phe Glu Thr Phe
                1160                1165                1170
Asn Trp Ala Leu Ser Met Gly Gly Lys Val Pro Val Ser Glu Gly
                1175                1180                1185
Leu Glu His Ser Asp Leu Pro Asp Gly Thr Gly Glu Phe Leu Asp
                1190                1195                1200
Ala Trp Leu Met Leu Val Glu Lys Met Val Asn Pro Thr Thr Val
                1205                1210                1215
Leu Glu Ser Pro His Ser Leu Pro Ala Lys Leu Pro Gly Gly Val
                1220                1225                1230
Gln Asn Phe Pro Gln Phe Ser Ala Leu Arg Phe Leu Val Val Thr
                1235                1240                1245
Gln Lys Ala Ala Phe Thr Cys Ile Lys Asn Leu Trp Asn Arg Lys
                1250                1255                1260
Pro Leu Lys Val Tyr Gly Gly Arg Met Ala Glu Ser Met Leu Ala
                1265                1270                1275
Ile Leu Cys His Ile Leu Arg Gly Glu Pro Val Ile Arg Glu Arg
                1280                1285                1290
Leu Ser Lys Glu Lys Glu Gly Ser Arg Gly Glu Glu Asp Thr Gly
                1295                1300                1305
Gln Glu Glu Gly Gly Ser Arg Arg Glu Pro Gln Val Asn Gln Gln
                1310                1315                1320
Gln Leu Gln Gln Leu Met Asp Met Gly Phe Thr Arg Glu His Ala
                1325                1330                1335
Met Glu Ala Leu Leu Asn Thr Ser Thr Met Glu Gln Ala Thr Glu
                1340                1345                1350
Tyr Leu Leu Thr His Pro Pro Pro Ile Met Gly Gly Val Val Arg
                1355                1360                1365
Asp Leu Ser Met Ser Glu Glu Asp Gln Met Met Arg Ala Ile Ala
                1370                1375                1380
Met Ser Leu Gly Gln Asp Ile Pro Met Asp Gln Arg Ala Glu Ser
                1385                1390                1395
Pro Glu Glu Val Ala Cys Arg Lys Glu Glu Glu Glu Arg Lys Ala
                1400                1405                1410
Arg Glu Lys Gln Glu Glu Glu Glu Ala Lys Cys Leu Glu Lys Phe
                1415                1420                1425
Gln Asp Ala Asp Pro Leu Glu Gln Asp Glu Leu His Thr Phe Thr
                1430                1435                1440
Asp Thr Met Leu Pro Gly Cys Phe His Leu Leu Asp Glu Leu Pro
                1445                1450                1455
Asp Thr Val Tyr Arg Val Cys Asp Leu Ile Met Thr Ala Ile Lys
                1460                1465                1470
Arg Asn Gly Ala Asp Tyr Arg Asp Met Ile Leu Lys Gln Val Val
                1475                1480                1485
Asn Gln Val Trp Glu Ala Ala Asp Val Leu Ile Lys Ala Ala Leu
                1490                1495                1500
Pro Leu Thr Thr Ser Asp Thr Lys Thr Val Ser Glu Trp Ile Ser
                1505                1510                1515
Gln Met Ala Thr Leu Pro Gln Ala Ser Asn Leu Ala Thr Arg Ile
                1520                1525                1530
Leu Leu Leu Thr Leu Leu Phe Glu Glu Leu Lys Leu Pro Cys Ala
```

1535 1540 1545
Trp Val Val Glu Ser Ser Gly Ile Leu Asn Val Leu Ile Lys Leu
1550 1555 1560
Leu Glu Val Val Gln Pro Cys Leu Gln Ala Ala Lys Glu Gln Lys
1565 1570 1575
Glu Val Gln Thr Pro Lys Trp Ile Thr Pro Val Leu Leu Leu Ile
1580 1585 1590
Asp Phe Tyr Glu Lys Thr Ala Ile Ser Ser Lys Arg Arg Ala Gln
1595 1600 1605
Met Thr Lys Tyr Leu Gln Ser Asn Ser Asn Asn Trp Arg Trp Phe
1610 1615 1620
Asp Asp Arg Ser Gly Arg Trp Cys Ser Tyr Ser Ala Ser Asn Asn
1625 1630 1635
Ser Thr Ile Asp Ser Ala Trp Lys Ser Gly Glu Thr Ser Val Arg
1640 1645 1650
Phe Thr Ala Gly Arg Arg Arg Tyr Thr Val Gln Phe Thr Thr Met
1655 1660 1665
Val Gln Val Asn Glu Glu Thr Gly Asn Arg Arg Pro Val Met Leu
1670 1675 1680
Thr Leu Leu Arg Val Pro Arg Leu Asn Lys Asn Ser Lys Asn Ser
1685 1690 1695
Asn Gly Gln Glu Leu Glu Lys Thr Leu Glu Glu Ser Lys Glu Met
1700 1705 1710
Asp Ile Lys Arg Lys Glu Asn Lys Gly Asn Asp Thr Pro Leu Ala
1715 1720 1725
Leu Glu Ser Thr Asn Thr Glu Lys Glu Thr Ser Leu Glu Glu Thr
1730 1735 1740
Lys Ile Gly Glu Ile Leu Ile Gln Gly Leu Thr Glu Asp Met Val
1745 1750 1755
Thr Val Leu Ile Arg Ala Cys Val Ser Met Leu Gly Val Pro Val
1760 1765 1770
Asp Pro Asp Thr Leu His Ala Thr Leu Arg Leu Cys Leu Arg Leu
1775 1780 1785
Thr Arg Asp His Lys Tyr Ala Met Met Phe Ala Glu Leu Lys Ser
1790 1795 1800
Thr Arg Met Ile Leu Asn Leu Thr Gln Ser Ser Gly Phe Asn Gly
1805 1810 1815
Phe Thr Pro Leu Val Thr Leu Leu Leu Arg His Ile Ile Glu Asp
1820 1825 1830
Pro Cys Thr Leu Arg His Thr Met Glu Lys Val Val Arg Ser Ala
1835 1840 1845
Ala Thr Ser Gly Ala Gly Ser Thr Thr Ser Gly Val Val Ser Gly
1850 1855 1860
Ser Leu Gly Ser Arg Glu Ile Asn Tyr Ile Leu Arg Val Leu Gly
1865 1870 1875
Pro Ala Ala Cys Arg Asn Pro Asp Ile Phe Thr Glu Val Ala Asn
1880 1885 1890
Cys Cys Ile Arg Ile Ala Leu Pro Ala Pro Arg Gly Ser Gly Thr
1895 1900 1905
Ala Ser Asp Asp Glu Phe Glu Asn Leu Arg Ile Lys Gly Pro Asn
1910 1915 1920
Ala Val Gln Leu Val Lys Thr Thr Pro Leu Lys Pro Ser Pro Leu
1925 1930 1935
Pro Val Ile Pro Asp Thr Ile Lys Glu Val Ile Tyr Asp Met Leu
1940 1945 1950
Asn Ala Leu Ala Ala Tyr His Ala Pro Glu Glu Ala Asp Lys Ser
1955 1960 1965
Asp Pro Lys Pro Gly Val Met Thr Gln Glu Val Gly Gln Leu Leu
1970 1975 1980
Gln Asp Met Gly Asp Asp Val Tyr Gln Gln Tyr Arg Ser Leu Thr
1985 1990 1995
Arg Gln Ser Ser Asp Phe Asp Thr Gln Ser Gly Phe Ser Ile Asn
2000 2005 2010

Ser Gln Val Phe Ala Ala Asp Gly Ala Ser Thr Glu Thr Ser Ala
2015 2020 2025
Ser Gly Thr Ser Gln Gly Glu Ala Ser Thr Pro Glu Glu Ser Arg
2030 2035 2040
Asp Gly Lys Lys Asp Lys Glu Gly Asp Arg Ala Ser Glu Glu Gly
2045 2050 2055
Lys Gln Lys Gly Lys Gly Ser Lys Pro Leu Met Pro Thr Ser Thr
2060 2065 2070
Ile Leu Arg Leu Leu Ala Glu Leu Val Arg Ser Tyr Val Gly Ile
2075 2080 2085
Ala Thr Leu Ile Ala Asn Tyr Ser Tyr Thr Val Gly Gln Ser Glu
2090 2095 2100
Leu Ile Lys Glu Asp Cys Ser Val Leu Ala Phe Val Leu Asp His
2105 2110 2115
Leu Leu Pro His Thr Gln Asn Ala Glu Asp Lys Asp Thr Pro Ala
2120 2125 2130
Leu Ala Arg Leu Phe Leu Ala Ser Leu Ala Ala Gly Ser Gly
2135 2140 2145
Thr Asp Ala Gln Val Ala Leu Val Asn Glu Val Lys Ala Ala Leu
2150 2155 2160
Gly Arg Ala Leu Ala Met Ala Glu Ser Thr Glu Lys His Ala Arg
2165 2170 2175
Leu Gln Ala Val Met Cys Ile Ile Ser Thr Ile Met Glu Ser Cys
2180 2185 2190
Pro Ser Thr Ser Ser Phe Tyr Ser Ser Ala Thr Ala Lys Thr Gln
2195 2200 2205
His Asn Gly Met Asn Asn Ile Ile Arg Leu Phe Leu Lys Lys Gly
2210 2215 2220
Leu Val Asn Asp Leu Ala Arg Val Pro His Ser Leu Asp Leu Ser
2225 2230 2235
Ser Pro Asn Met Ala Asn Thr Val Asn Ala Ala Leu Lys Pro Leu
2240 2245 2250
Glu Thr Leu Ser Arg Ile Val Asn Gln Pro Ser Ser Leu Phe Gly
2255 2260 2265
Ser Lys Ser Ala Ser Ser Lys Asn Lys Ser Glu Gln Asp Ala Gln
2270 2275 2280
Gly Ala Ser Gln Asp Ser Ser Ser Asn Gln Gln Asp Pro Gly Glu
2285 2290 2295
Pro Gly Glu Ala Glu Val Gln Glu Glu Asp His Asp Val Thr Gln
2300 2305 2310
Thr Glu Val Ala Asp Gly Asp Ile Met Asp Gly Glu Ala Glu Thr
2315 2320 2325
Asp Ser Val Val Ile Ala Gly Gln Pro Glu Val Leu Ser Ser Gln
2330 2335 2340
Glu Met Gln Val Glu Asn Glu Leu Glu Asp Leu Ile Asp Glu Leu
2345 2350 2355
Leu Glu Arg Asp Gly Gly Ser Gly Asn Ser Thr Ile Ile Val Ser
2360 2365 2370
Arg Ser Gly Glu Asp Glu Ser Gln Glu Asp Val Leu Met Asp Glu
2375 2380 2385
Ala Pro Ser Asn Leu Ser Gln Ala Ser Thr Leu Gln Ala Asn Arg
2390 2395 2400
Glu Asp Ser Met Asn Ile Leu Asp Pro Glu Asp Glu Glu Glu His
2405 2410 2415
Thr Gln Glu Glu Asp Ser Ser Gly Ser Asn Glu Asp Glu Asp Asp
2420 2425 2430
Ser Gln Asp Glu Glu Glu Glu Glu Glu Glu Asp Glu Glu Asp Asp
2435 2440 2445
Gln Glu Asp Asp Glu Gly Glu Glu Gly Asp Glu Asp Asp Asp Asp
2450 2455 2460
Asp Gly Ser Glu Met Glu Leu Asp Glu Asp Tyr Pro Asp Met Asn
2465 2470 2475
Ala Ser Pro Leu Val Arg Phe Glu Arg Phe Asp Arg Glu Asp Asp

2480 2485 2490
Leu Ile Ile Glu Phe Asp Asn Met Phe Ser Ser Ala Thr Asp Ile
2495 2500 2505
Pro Pro Ser Pro Gly Asn Ile Pro Thr Thr His Pro Leu Met Val
2510 2515 2520
Arg His Ala Asp His Ser Ser Leu Thr Leu Gly Ser Gly Ser Ser
2525 2530 2535
Thr Thr Arg Leu Thr Gln Gly Ile Gly Arg Ser Gln Arg Thr Leu
2540 2545 2550
Arg Gln Leu Thr Ala Asn Thr Gly His Thr Ile His Val His Tyr
2555 2560 2565
Pro Gly Asn Arg Gln Pro Asn Pro Pro Leu Ile Leu Gln Arg Leu
2570 2575 2580
Leu Gly Pro Ser Ala Ala Ala Asp Ile Leu Gln Leu Ser Ser Ser
2585 2590 2595
Leu Pro Leu Gln Ser Arg Gly Arg Ala Arg Leu Leu Val Gly Asn
2600 2605 2610
Asp Asp Val His Ile Ile Ala Arg Ser Asp Asp Glu Leu Leu Asp
2615 2620 2625
Asp Phe Phe His Asp Gln Ser Thr Ala Thr Ser Gln Ala Gly Thr
2630 2635 2640
Leu Ser Ser Ile Pro Thr Ala Leu Thr Arg Trp Thr Glu Glu Cys
2645 2650 2655
Lys Val Leu Asp Ala Glu Ser Met His Asp Cys Val Ser Val Val
2660 2665 2670
Lys Val Ser Ile Val Asn His Leu Glu Phe Leu Arg Asp Glu Glu
2675 2680 2685
Leu Glu Glu Arg Arg Glu Lys Arg Arg Lys Gln Leu Ala Glu Glu
2690 2695 2700
Glu Thr Lys Ile Thr Asp Lys Gly Lys Glu Asp Lys Glu Asn Arg
2705 2710 2715
Asp Gln Ser Ala Gln Cys Thr Ala Ser Lys Ser Asn Asp Ser Thr
2720 2725 2730
Glu Gln Asn Leu Ser Asp Gly Thr Pro Met Pro Asp Ser Tyr Pro
2735 2740 2745
Thr Thr Pro Ser Ser Thr Asp Ala Ala Thr Ser Glu Ser Lys Glu
2750 2755 2760
Thr Leu Gly Thr Leu Gln Ser Ser Gln Gln Gln Pro Thr Leu Pro
2765 2770 2775
Thr Pro Pro Ala Leu Gly Glu Val Pro Gln Glu Leu Gln Ser Pro
2780 2785 2790
Ala Gly Glu Gly Gly Ser Ser Thr Gln Leu Leu Met Pro Val Glu
2795 2800 2805
Pro Glu Glu Leu Gly Pro Thr Arg Pro Ser Gly Glu Ala Glu Thr
2810 2815 2820
Thr Gln Met Glu Leu Ser Pro Ala Pro Thr Ile Thr Ser Leu Ser
2825 2830 2835
Pro Glu Arg Ala Glu Asp Ser Asp Ala Leu Thr Ala Val Ser Ser
2840 2845 2850
Gln Leu Glu Gly Ser Pro Met Asp Thr Ser Ser Leu Ala Ser Cys
2855 2860 2865
Thr Leu Glu Glu Ala Val Gly Asp Thr Ser Ala Ala Gly Ser Ser
2870 2875 2880
Glu Gln Pro Arg Ala Gly Ser Ser Thr Pro Gly Asp Ala Pro Pro
2885 2890 2895
Ala Val Ala Glu Val Gln Gly Arg Ser Asp Gly Ser Gly Glu Ser
2900 2905 2910
Ala Gln Pro Pro Glu Asp Ser Ser Pro Pro Ala Ser Ser Glu Ser
2915 2920 2925
Ser Ser Thr Arg Asp Ser Ala Val Ala Ile Ser Gly Ala Asp Ser
2930 2935 2940
Arg Gly Ile Leu Glu Glu Pro Leu Pro Ser Thr Ser Ser Glu Glu
2945 2950 2955

Glu Asp Pro Leu Ala Gly Ile Ser Leu Pro Glu Gly Val Asp Pro
2960 2965 2970
Ser Phe Leu Ala Ala Leu Pro Asp Asp Ile Arg Arg Glu Val Leu
2975 2980 2985
Gln Asn Gln Leu Gly Ile Arg Pro Pro Thr Arg Thr Ala Pro Ser
2990 2995 3000
Thr Asn Ser Ser Ala Pro Ala Val Val Gly Asn Pro Gly Val Thr
3005 3010 3015
Glu Val Ser Pro Glu Phe Leu Ala Ala Leu Pro Pro Ala Ile Gln
3020 3025 3030
Glu Glu Val Leu Ala Gln Gln Arg Ala Glu Gln Gln Arg Arg Glu
3035 3040 3045
Leu Ala Gln Asn Ala Ser Ser Asp Thr Pro Met Asp Pro Val Thr
3050 3055 3060
Phe Ile Gln Thr Leu Pro Ser Asp Leu Arg Arg Ser Val Leu Glu
3065 3070 3075
Asp Met Glu Asp Ser Val Leu Ala Val Met Pro Pro Asp Ile Ala
3080 3085 3090
Ala Glu Ala Gln Ala Leu Arg Arg Glu Gln Glu Ala Arg Gln Arg
3095 3100 3105
Gln Leu Met His Glu Arg Leu Phe Gly His Ser Ser Thr Ser Ala
3110 3115 3120
Leu Ser Ala Ile Leu Arg Ser Pro Ala Phe Thr Ser Arg Leu Ser
3125 3130 3135
Gly Asn Arg Gly Val Gln Tyr Thr Arg Leu Ala Val Gln Arg Gly
3140 3145 3150
Gly Thr Phe Gln Met Gly Gly Ser Ser Ser His Asn Arg Pro Ser
3155 3160 3165
Gly Ser Asn Val Asp Thr Leu Leu Arg Leu Arg Gly Arg Leu Leu
3170 3175 3180
Leu Asp His Glu Ala Leu Ser Cys Leu Leu Val Leu Leu Phe Val
3185 3190 3195
Asp Glu Pro Lys Leu Asn Thr Ser Arg Leu His Arg Val Leu Arg
3200 3205 3210
Asn Leu Cys Tyr His Ala Gln Thr Arg His Trp Val Ile Arg Ser
3215 3220 3225
Leu Leu Ser Ile Leu Gln Arg Ser Ser Glu Ser Glu Leu Cys Ile
3230 3235 3240
Glu Thr Pro Lys Leu Thr Thr Ser Glu Glu Lys Gly Lys Lys Ser
3245 3250 3255
Ser Lys Ser Cys Gly Ser Ser Ser His Glu Asn Arg Pro Leu Asp
3260 3265 3270
Leu Leu His Lys Met Glu Ser Lys Ser Ser Asn Gln Leu Ser Trp
3275 3280 3285
Leu Ser Val Ser Met Asp Ala Ala Leu Gly Cys Arg Thr Asn Ile
3290 3295 3300
Phe Gln Ile Gln Arg Ser Gly Gly Arg Lys His Thr Glu Lys His
3305 3310 3315
Ala Ser Gly Gly Ser Thr Val His Ile His Pro Gln Ala Ala Pro
3320 3325 3330
Val Val Cys Arg His Val Leu Asp Thr Leu Ile Gln Leu Ala Lys
3335 3340 3345
Val Phe Pro Ser His Phe Thr Gln Gln Arg Thr Lys Glu Thr Asn
3350 3355 3360
Cys Glu Ser Asp Arg Glu Arg Gly Asn Lys Ala Cys Ser Pro Cys
3365 3370 3375
Ser Ser Gln Ser Ser Ser Ser Gly Ile Cys Thr Asp Phe Trp Asp
3380 3385 3390
Leu Leu Val Lys Leu Asp Asn Met Asn Val Ser Arg Lys Gly Lys
3395 3400 3405
Asn Ser Val Lys Ser Val Pro Val Ser Ala Gly Gly Glu Gly Glu
3410 3415 3420
Thr Ser Pro Tyr Ser Leu Glu Ala Ser Pro Leu Gly Gln Leu Met

```
                3425                    3430                    3435
Asn Met Leu Ser His Pro Val Ile Arg Arg Ser Ser Leu Leu Thr
                3440                    3445                    3450
Glu Lys Leu Leu Arg Leu Leu Ser Leu Ile Ser Ile Ala Leu Pro
                3455                    3460                    3465
Glu Asn Lys Val Ser Glu Ala Gln Ala Asn Ser Gly Ser Gly Ala
                3470                    3475                    3480
Ser Ser Thr Thr Thr Ala Thr Ser Thr Thr Ser Thr Thr Thr Thr
                3485                    3490                    3495
Thr Ala Ala Ser Thr Thr Pro Thr Pro Pro Thr Ala Pro Thr Pro
                3500                    3505                    3510
Val Thr Ser Ala Pro Ala Leu Val Ala Ala Thr Ala Ile Ser Thr
                3515                    3520                    3525
Ile Val Val Ala Ala Ser Thr Thr Val Thr Thr Pro Thr Thr Ala
                3530                    3535                    3540
Thr Thr Thr Val Ser Ile Ser Pro Thr Thr Lys Gly Ser Lys Ser
                3545                    3550                    3555
Pro Ala Lys Val Ser Asp Gly Gly Ser Ser Ser Thr Asp Phe Lys
                3560                    3565                    3570
Met Val Ser Ser Gly Leu Thr Glu Asn Gln Leu Gln Leu Ser Val
                3575                    3580                    3585
Glu Val Leu Thr Ser His Ser Cys Ser Glu Glu Gly Leu Glu Asp
                3590                    3595                    3600
Ala Ala Asn Val Leu Leu Gln Leu Ser Arg Gly Asp Ser Gly Thr
                3605                    3610                    3615
Arg Asp Thr Val Leu Lys Leu Leu Leu Asn Gly Ala Arg His Leu
                3620                    3625                    3630
Gly Tyr Thr Leu Cys Lys Gln Ile Gly Thr Leu Leu Ala Glu Leu
                3635                    3640                    3645
Arg Glu Tyr Asn Leu Glu Gln Gln Arg Arg Ala Gln Cys Glu Thr
                3650                    3655                    3660
Leu Ser Pro Asp Gly Leu Pro Glu Glu Gln Pro Gln Thr Thr Lys
                3665                    3670                    3675
Leu Lys Gly Lys Met Gln Ser Arg Phe Asp Met Ala Glu Asn Val
                3680                    3685                    3690
Val Ile Val Ala Ser Gln Lys Arg Pro Leu Gly Gly Arg Glu Leu
                3695                    3700                    3705
Gln Leu Pro Ser Met Ser Met Leu Thr Ser Lys Thr Ser Thr Gln
                3710                    3715                    3720
Lys Phe Phe Leu Arg Val Leu Gln Val Ile Ile Gln Leu Arg Asp
                3725                    3730                    3735
Asp Thr Arg Arg Ala Asn Lys Lys Ala Lys Gln Thr Gly Arg Leu
                3740                    3745                    3750
Gly Ser Ser Gly Leu Gly Ser Ala Ser Ser Ile Gln Ala Ala Val
                3755                    3760                    3765
Arg Gln Leu Glu Ala Glu Ala Asp Ala Ile Ile Gln Met Val Arg
                3770                    3775                    3780
Glu Gly Gln Arg Ala Arg Arg Gln Gln Gln Ala Ala Thr Ser Glu
                3785                    3790                    3795
Ser Ser Gln Ser Glu Ala Ser Val Arg Arg Glu Glu Ser Pro Met
                3800                    3805                    3810
Asp Val Asp Gln Pro Ser Pro Ser Ala Gln Asp Thr Gln Ser Ile
                3815                    3820                    3825
Ala Ser Asp Gly Thr Pro Gln Gly Glu Lys Glu Lys Glu Glu Arg
                3830                    3835                    3840
Pro Pro Glu Leu Pro Leu Leu Ser Glu Gln Leu Ser Leu Asp Glu
                3845                    3850                    3855
Leu Trp Asp Met Leu Gly Glu Cys Leu Lys Glu Leu Glu Glu Ser
                3860                    3865                    3870
His Asp Gln His Ala Val Leu Val Leu Gln Pro Ala Val Glu Ala
                3875                    3880                    3885
Phe Phe Leu Val His Ala Thr Glu Arg Glu Ser Lys Pro Pro Val
                3890                    3895                    3900
```

Arg Asp Thr Arg Glu Ser Gln Leu Ala His Ile Lys Asp Glu Pro
3905 3910 3915
Pro Pro Leu Ser Pro Ala Pro Leu Thr Pro Ala Thr Pro Ser Ser
3920 3925 3930
Leu Asp Pro Phe Phe Ser Arg Glu Pro Ser Ser Met His Ile Ser
3935 3940 3945
Ser Ser Leu Pro Pro Asp Thr Gln Lys Phe Leu Arg Phe Ala Glu
3950 3955 3960
Thr His Arg Thr Val Leu Asn Gln Ile Leu Arg Gln Ser Thr Thr
3965 3970 3975
His Leu Ala Asp Gly Pro Phe Ala Val Leu Val Asp Tyr Ile Arg
3980 3985 3990
Val Leu Asp Phe Asp Val Lys Arg Lys Tyr Phe Arg Gln Glu Leu
3995 4000 4005
Glu Arg Leu Asp Glu Gly Leu Arg Lys Glu Asp Met Ala Val His
4010 4015 4020
Val Arg Arg Asp His Val Phe Glu Asp Ser Tyr Arg Glu Leu His
4025 4030 4035
Arg Lys Ser Pro Glu Glu Met Lys Asn Arg Leu Tyr Ile Val Phe
4040 4045 4050
Glu Gly Glu Glu Gly Gln Asp Ala Gly Gly Leu Leu Arg Glu Trp
4055 4060 4065
Tyr Met Ile Ile Ser Arg Glu Met Phe Asn Pro Met Tyr Ala Leu
4070 4075 4080
Phe Arg Thr Ser Pro Gly Asp Arg Val Thr Tyr Thr Ile Asn Pro
4085 4090 4095
Ser Ser His Cys Asn Pro Asn His Leu Ser Tyr Phe Lys Phe Val
4100 4105 4110
Gly Arg Ile Val Ala Lys Ala Val Tyr Asp Asn Arg Leu Leu Glu
4115 4120 4125
Cys Tyr Phe Thr Arg Ser Phe Tyr Lys His Ile Leu Gly Lys Ser
4130 4135 4140
Val Arg Tyr Thr Asp Met Glu Ser Glu Asp Tyr His Phe Tyr Gln
4145 4150 4155
Gly Leu Val Tyr Leu Leu Glu Asn Asp Val Ser Thr Leu Gly Tyr
4160 4165 4170
Asp Leu Thr Phe Ser Thr Glu Val Gln Glu Phe Gly Val Cys Glu
4175 4180 4185
Val Arg Asp Leu Lys Pro Asn Gly Ala Asn Ile Leu Val Thr Glu
4190 4195 4200
Glu Asn Lys Lys Glu Tyr Val His Leu Val Cys Gln Met Arg Met
4205 4210 4215
Thr Gly Ala Ile Arg Lys Gln Leu Ala Ala Phe Leu Glu Gly Phe
4220 4225 4230
Tyr Glu Ile Ile Pro Lys Arg Leu Ile Ser Ile Phe Thr Glu Gln
4235 4240 4245
Glu Leu Glu Leu Leu Ile Ser Gly Leu Pro Thr Ile Asp Ile Asp
4250 4255 4260
Asp Leu Lys Ser Asn Thr Glu Tyr His Lys Tyr Gln Ser Asn Ser
4265 4270 4275
Ile Gln Ile Gln Trp Phe Trp Arg Ala Leu Arg Ser Phe Asp Gln
4280 4285 4290
Ala Asp Arg Ala Lys Phe Leu Gln Phe Val Thr Gly Thr Ser Lys
4295 4300 4305
Val Pro Leu Gln Gly Phe Ala Ala Leu Glu Gly Met Asn Gly Ile
4310 4315 4320
Gln Lys Phe Gln Ile His Arg Asp Asp Arg Ser Thr Asp Arg Leu
4325 4330 4335
Pro Ser Ala His Thr Cys Phe Asn Gln Leu Asp Leu Pro Ala Tyr
4340 4345 4350
Glu Ser Phe Glu Lys Leu Arg His Met Leu Leu Leu Ala Ile Gln
4355 4360 4365
Glu Cys Ser Glu Gly Phe Gly Leu Ala

4370

<210> 36
<211> 89
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504732CD1

<400> 36
Met Ala Ala Trp Lys Ser Trp Thr Ala Leu Arg Leu Cys Ala Thr
1 5 10 15
Val Val Val Leu Asp Met Val Val Cys Lys Gly Phe Val Glu Asp
20 25 30
Leu Asp Glu Ser Phe Lys Glu Asn Arg Asn Asp Asp Ile Trp Leu
35 40 45
Val Asp Phe Tyr Ala Pro Trp Cys Gly His Cys Lys Lys Leu Glu
50 55 60
Pro Ile Trp Asn Glu Val Gly Leu Glu Met Lys Ser Ile Gly Ser
65 70 75
Pro Val Lys Val Gly Lys Met Asp Ala Thr Ser Tyr Ser Asn
80 85

<210> 37
<211> 424
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 950917CD1

<400> 37
Met Pro Pro Phe Leu Ile Thr Leu Phe Leu Phe His Ser Cys Cys
1 5 10 15
Leu Arg Ala Asn Gly His Leu Arg Glu Gly Met Thr Leu Leu Lys
20 25 30
Thr Glu Phe Ala Leu His Leu Tyr Gln Ser Val Ala Ala Cys Arg
35 40 45
Asn Glu Thr Asn Phe Val Ile Ser Pro Ala Gly Val Ser Leu Pro
50 55 60
Leu Glu Ile Leu Gln Phe Gly Ala Glu Gly Ser Thr Gly Gln Gln
65 70 75
Leu Ala Asp Ala Leu Gly Tyr Thr Val His Asp Lys Arg Val Lys
80 85 90
Asp Phe Leu His Ala Val Tyr Ala Thr Leu Pro Thr Ser Ser Gln
95 100 105
Gly Thr Glu Met Glu Leu Ala Cys Ser Leu Phe Val Gln Val Gly
110 115 120
Thr Pro Leu Ser Pro Cys Phe Val Glu His Val Ser Trp Trp Ala
125 130 135
Asn Ser Ser Leu Glu Pro Ala Asp Leu Ser Glu Pro Asn Ser Thr
140 145 150
Ala Ile Gln Thr Ser Glu Gly Ala Ser Arg Glu Thr Ala Gly Gly
155 160 165
Gly Pro Ser Glu Gly Pro Gly Gly Trp Pro Trp Glu Gln Val Ser
170 175 180
Ala Ala Phe Ala Gln Leu Val Leu Val Ser Thr Met Ser Phe Gln
185 190 195
Gly Thr Trp Arg Lys Arg Phe Ser Ser Thr Asp Thr Gln Ile Leu
200 205 210

Pro Phe Thr Cys Ala Tyr Gly Leu Val Leu Gln Val Pro Met Met
215 220 225
His Gln Thr Thr Glu Val Asn Tyr Gly Gln Phe Gln Asp Thr Ala
230 235 240
Gly His Gln Val Gly Val Leu Glu Leu Pro Tyr Leu Gly Ser Ala
245 250 255
Val Ser Leu Phe Leu Val Leu Pro Arg Asp Lys Asp Thr Pro Leu
260 265 270
Ser His Ile Glu Pro His Leu Thr Ala Ser Thr Ile His Leu Trp
275 280 285
Thr Thr Ser Leu Arg Arg Ala Arg Met Asp Val Phe Leu Pro Arg
290 295 300
Phe Arg Ile Gln Asn Gln Phe Asn Leu Lys Ser Ile Leu Asn Ser
305 310 315
Trp Gly Val Thr Asp Leu Phe Asp Pro Leu Lys Ala Asn Leu Lys
320 325 330
Gly Ile Ser Gly Gln Asp Gly Phe Tyr Val Ser Glu Ala Ile His
335 340 345
Lys Ala Lys Ile Glu Val Leu Glu Glu Gly Thr Lys Ala Ser Gly
350 355 360
Ala Thr Ala Leu Leu Leu Leu Lys Arg Ser Arg Ile Pro Ile Phe
365 370 375
Lys Ala Asp Arg Pro Phe Ile Tyr Phe Leu Arg Glu Pro Asn Thr
380 385 390
Gly Ile Thr Val Phe Phe Asp Arg Ile Gln Ile Ile Tyr Gln Cys
395 400 405
Leu Ser Ser Asn Lys Gly Ser Phe Val His Tyr Pro Leu Lys Asn
410 415 420
Lys His Ser Phe

<210> 38
<211> 791
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7459720CD1

<400> 38
Met Pro Lys Gly Arg Gln Lys Val Pro His Leu Asp Ala Pro Leu
1 5 10 15
Gly Leu Pro Thr Cys Leu Trp Leu Glu Leu Ala Gly Leu Phe Leu
20 25 30
Leu Val Pro Trp Val Met Gly Leu Ala Gly Thr Gly Gly Pro Asp
35 40 45
Gly Gln Gly Thr Gly Gly Pro Ser Trp Ala Val His Leu Glu Ser
50 55 60
Leu Glu Gly Asp Gly Glu Glu Glu Thr Leu Glu Gln Gln Ala Asp
65 70 75
Ala Leu Ala Gln Ala Ala Gly Leu Val Asn Ala Gly Arg Ile Gly
80 85 90
Glu Leu Gln Gly His Tyr Leu Phe Val Gln Pro Ala Gly His Arg
95 100 105
Pro Ala Leu Glu Val Glu Ala Ile Arg Gln Gln Val Glu Ala Val
110 115 120
Leu Ala Gly His Glu Ala Val Arg Trp His Ser Glu Gln Arg Leu
125 130 135
Leu Arg Arg Ala Lys Arg Ser Val His Phe Asn Asp Pro Lys Tyr
140 145 150
Pro Gln Gln Trp His Leu Asn Asn Arg Arg Ser Pro Gly Arg Asp
155 160 165

Ile Asn Val Thr Gly Val Trp Glu Arg Asn Val Thr Gly Arg Gly
170 175 180
Val Thr Val Val Val Val Asp Asp Gly Val Glu His Thr Ile Gln
185 190 195
Asp Ile Ala Pro Asn Tyr Ser Pro Glu Gly Ser Tyr Asp Leu Asn
200 205 210
Ser Asn Asp Pro Asp Pro Met Pro His Pro Asp Val Glu Asn Gly
215 220 225
Asn His His Gly Thr Arg Cys Ala Gly Glu Ile Ala Ala Val Pro
230 235 240
Asn Asn Ser Phe Cys Ala Val Gly Val Ala Tyr Gly Ser Arg Ile
245 250 255
Ala Gly Ile Arg Val Leu Asp Gly Pro Leu Thr Asp Ser Met Glu
260 265 270
Ala Val Ala Phe Asn Lys His Tyr Gln Ile Asn Asp Ile Tyr Ser
275 280 285
Cys Arg Tyr Leu Phe Leu Asn Ser Trp Gly Pro Asp Asp Asp Gly
290 295 300
Lys Thr Val Asp Gly Pro His Gln Leu Gly Lys Ala Ala Leu Gln
305 310 315
His Gly Val Ile Ala Gly Arg Gln Gly Phe Gly Ser Ile Phe Val
320 325 330
Val Ala Ser Gly Asn Gly Gly Gln His Asn Asp Asn Cys Asn Tyr
335 340 345
Asp Gly Tyr Ala Asn Ser Ile Tyr Thr Val Thr Ile Gly Ala Val
350 355 360
Asp Glu Glu Gly Arg Met Pro Phe Tyr Ala Glu Glu Cys Ala Ser
365 370 375
Met Leu Ala Val Thr Phe Ser Gly Gly Asp Lys Met Leu Arg Ser
380 385 390
Ile Val Thr Thr Asp Trp Asp Leu Gln Lys Gly Thr Gly Cys Thr
395 400 405
Glu Gly His Thr Gly Thr Ser Ala Ala Ala Pro Leu Ala Ala Gly
410 415 420
Met Ile Ala Leu Met Leu Gln Val Arg Pro Cys Leu Thr Trp Arg
425 430 435
Asp Val Gln His Ile Ile Val Phe Thr Ala Thr Arg Tyr Glu Asp
440 445 450
Arg Arg Ala Glu Trp Val Thr Asn Glu Ala Gly Phe Ser His Ser
455 460 465
His Gln His Gly Phe Gly Leu Leu Asn Ala Trp Arg Leu Val Asn
470 475 480
Ala Ala Lys Ile Trp Thr Ser Val Pro Tyr Leu Ala Ser Tyr Val
485 490 495
Ser Pro Val Leu Lys Glu Asn Lys Ala Ile Pro Gln Ser Pro Arg
500 505 510
Ser Leu Glu Val Leu Trp Asn Val Ser Arg Met Asp Leu Glu Met
515 520 525
Ser Gly Leu Lys Thr Leu Glu His Val Ala Val Thr Val Ser Ile
530 535 540
Thr His Pro Arg Arg Gly Ser Leu Glu Leu Lys Leu Phe Cys Pro
545 550 555
Ser Gly Met Met Ser Leu Ile Gly Ala Pro Arg Ser Met Asp Ser
560 565 570
Asp Pro Asn Gly Phe Asn Asp Trp Thr Phe Ser Thr Val Arg Cys
575 580 585
Trp Gly Glu Arg Ala Arg Gly Thr Tyr Arg Leu Val Ile Arg Asp
590 595 600
Val Gly Asp Glu Ser Phe Gln Val Gly Ile Leu Arg Gln Trp Gln
605 610 615
Leu Thr Leu Tyr Gly Ser Val Trp Ser Ala Val Asp Ile Arg Asp
620 625 630
Arg Gln Arg Leu Leu Glu Ser Ala Met Ser Gly Lys Tyr Leu His

```
                635                 640                 645
Asp Asp Phe Ala Leu Pro Cys Pro Pro Gly Leu Lys Ile Pro Glu
                650                 655                 660
Glu Asp Gly Tyr Thr Ile Thr Pro Asn Thr Leu Lys Thr Leu Val
                665                 670                 675
Leu Val Gly Cys Phe Thr Val Phe Trp Thr Val Tyr Tyr Met Leu
                680                 685                 690
Glu Val Tyr Leu Ser Gln Arg Asn Val Ala Ser Asn Gln Val Cys
                695                 700                 705
Arg Ser Gly Pro Cys His Trp Pro His Arg Ser Arg Lys Ala Lys
                710                 715                 720
Glu Glu Gly Thr Glu Leu Glu Ser Val Pro Leu Cys Ser Ser Lys
                725                 730                 735
Asp Pro Asp Glu Val Glu Thr Glu Ser Arg Gly Pro Pro Thr Thr
                740                 745                 750
Ser Asp Leu Leu Ala Pro Asp Leu Leu Glu Gln Gly Asp Trp Ser
                755                 760                 765
Leu Ser Gln Asn Lys Ser Ala Leu Asp Cys Pro His Gln His Leu
                770                 775                 780
Asp Val Pro His Gly Lys Glu Glu Gln Ile Cys
                785                 790
```

<210> 39
<211> 352
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503300CD1

<400> 39

```
Met Leu Cys Val Gly Arg Leu Gly Gly Leu Gly Ala Arg Ala Ala
 1               5                  10                  15
Ala Leu Pro Pro Arg Arg Ala Gly Arg Gly Ser Leu Glu Ala Gly
                20                  25                  30
Ile Arg Ala Arg Arg Val Ser Thr Ser Trp Ser Pro Val Gly Ala
                35                  40                  45
Ala Phe Asn Val Lys Pro Gln Gly Ser Arg Leu Asp Leu Phe Gly
                50                  55                  60
Glu Arg Arg Gly Leu Phe Gly Val Pro Glu Leu Ser Ala Pro Glu
                65                  70                  75
Gly Phe His Ile Ala Gln Glu Lys Ala Leu Arg Lys Thr Glu Leu
                80                  85                  90
Leu Val Asp Arg Ala Cys Ser Thr Pro Pro Gly Pro Gln Thr Val
                95                 100                 105
Leu Ile Phe Asp Glu Leu Ser Asp Ser Leu Cys Arg Val Ala Asp
               110                 115                 120
Leu Ala Asp Phe Val Lys Ile Ala His Pro Glu Pro Ala Phe Arg
               125                 130                 135
Glu Ala Ala Glu Glu Ala Cys Arg Ser Ile Gly Thr Met Val Glu
               140                 145                 150
Lys Leu Asn Thr Asn Val Asp Leu Tyr Gln Ser Leu Gln Lys Leu
               155                 160                 165
Leu Ala Asp Lys Lys Leu Val Asp Ser Leu Asp Pro Glu Thr Arg
               170                 175                 180
Arg Val Ala Glu Leu Phe Met Phe Asp Phe Glu Ile Ser Gly Ile
               185                 190                 195
His Leu Asp Lys Glu Lys Arg Lys Arg Ala Val Asp Leu Asn Val
               200                 205                 210
Lys Ile Leu Asp Leu Ser Ser Thr Phe Leu Met Gly Thr Asn Phe
               215                 220                 225
Pro Asn Lys Ile Glu Lys His Leu Leu Pro Glu His Ile Arg Arg
```

```
                    230                           235                           240
Asn Phe Thr Ser Ala Gly Asp His Ile Ile Ile Asp Gly Leu His
                    245                           250                           255
Ala Glu Ser Pro Asp Asp Leu Val Arg Glu Ala Ala Tyr Lys Ile
                    260                           265                           270
Phe Leu Tyr Pro Asn Ala Gly Gln Leu Lys Cys Leu Glu Glu Leu
                    275                           280                           285
Leu Ser Ser Arg Asp Leu Leu Ala Lys Leu Val Gly Tyr Ser Thr
                    290                           295                           300
Phe Ser His Arg Ala Leu Gln Gly Thr Ile Ala Lys Asn Pro Glu
                    305                           310                           315
Thr Val Met Gln Phe Leu Glu Lys Leu Ser Asp Lys Leu Ser Glu
                    320                           325                           330
Arg Thr Leu Lys Asp Phe Glu Met Ile Arg Gly Met Lys Met Lys
                    335                           340                           345
Leu Asn Pro Gln Asn Ser Val
                    350
```

<210> 40
<211> 495
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503334CD1

<400> 40

```
Met Glu Leu His Gly Gly Asn Arg Gly Gly Tyr Gln Pro Lys Ala
 1                   5                            10                            15
Leu Glu Gln Ser Gly Cys Trp Leu Trp Gln Ala Ala Leu Ser Thr
                    20                            25                            30
Leu Asn Pro Asn Pro Thr Asp Ser Cys Pro Leu Tyr Leu Asn Tyr
                    35                            40                            45
Ala Thr Val Ala Ala Leu Pro Cys Arg Val Ser Arg His Asn Ser
                    50                            55                            60
Pro Ser Ala Ala His Phe Ile Thr Arg Leu Val Arg Thr Cys Leu
                    65                            70                            75
Pro Pro Gly Ala His Arg Cys Ile Val Met Val Cys Glu Gln Pro
                    80                            85                            90
Glu Val Phe Ala Ser Ala Cys Ala Leu Ala Arg Ala Phe Pro Leu
                    95                            100                           105
Phe Thr His Arg Ser Gly Ala Ser Arg Arg Leu Glu Lys Lys Thr
                    110                           115                           120
Val Thr Val Glu Phe Phe Leu Val Gly Gln Asp Asn Gly Pro Val
                    125                           130                           135
Glu Val Ser Thr Leu Gln Cys Leu Ala Asn Ala Thr Asp Gly Val
                    140                           145                           150
Arg Leu Ala Ala Arg Ile Val Asp Thr Pro Cys Asn Glu Met Asn
                    155                           160                           165
Thr Asp Thr Phe Leu Glu Glu Ile Asn Lys Val Gly Lys Glu Leu
                    170                           175                           180
Gly Ile Ile Pro Thr Ile Ile Arg Asp Glu Glu Leu Lys Thr Arg
                    185                           190                           195
Gly Phe Gly Gly Ile Tyr Gly Val Gly Lys Ala Ala Leu His Pro
                    200                           205                           210
Pro Ala Leu Ala Val Leu Ser His Thr Pro Asp Gly Ala Thr Gln
                    215                           220                           225
Thr Ile Ala Trp Val Gly Lys Gly Ile Val Tyr Asp Thr Gly Gly
                    230                           235                           240
Leu Ser Ile Lys Gly Lys Thr Thr Met Pro Gly Met Lys Arg Asp
                    245                           250                           255
Cys Gly Gly Ala Ala Ala Val Leu Gly Ala Phe Arg Ala Ala Ile
```

260 265 270
Lys Gln Gly Phe Lys Asp Asn Leu His Ala Val Phe Cys Leu Ala
275 280 285
Glu Asn Ser Val Gly Pro Asn Ala Thr Arg Pro Asp Asp Ile His
290 295 300
Leu Leu Tyr Ser Gly Lys Thr Val Glu Ile Asn Asn Thr Asp Ala
305 310 315
Glu Gly Arg Leu Val Leu Ala Asp Gly Val Ser Tyr Ala Cys Lys
320 325 330
Asp Leu Gly Ala Asp Ile Ile Leu Asp Met Ala Thr Leu Thr Gly
335 340 345
Ala Gln Gly Ile Ala Thr Gly Lys Tyr His Ala Ala Val Leu Thr
350 355 360
Asn Ser Ala Glu Trp Glu Ala Ala Cys Val Lys Ala Gly Arg Lys
365 370 375
Cys Gly Asp Leu Val His Pro Leu Val Tyr Cys Pro Glu Leu His
380 385 390
Phe Ser Glu Phe Thr Ser Ala Val Ala Asp Met Lys Asn Ser Leu
395 400 405
Ala Asp Arg Asp Asn Ser Pro Ser Ser Cys Ala Gly Leu Phe Ile
410 415 420
Ala Ser His Ile Gly Phe Asp Trp Pro Gly Val Trp Val His Leu
425 430 435
Asp Ile Ala Ala Pro Val His Ala Gly Glu Arg Ala Thr Gly Phe
440 445 450
Gly Val Ala Leu Leu Leu Ala Leu Phe Gly Arg Ala Ser Glu Asp
455 460 465
Pro Leu Leu Asn Leu Val Ser Pro Leu Gly Cys Glu Val Asp Val
470 475 480
Glu Glu Gly Asp Leu Gly Arg Asp Ser Lys Arg Arg Arg Leu Val
485 490 495

<210> 41
<211> 239
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503341CD1

<400> 41
Met Ala Ala Ser Thr Gly Tyr Val Arg Leu Trp Gly Ala Ala Arg
1 5 10 15
Cys Trp Val Leu Arg Arg Pro Met Leu Ala Ala Ala Gly Gly Arg
20 25 30
Val Pro Thr Ala Ala Gly Ala Trp Leu Leu Arg Gly Gln Arg Thr
35 40 45
Cys Asp Ala Ser Pro Pro Trp Ala Leu Trp Gly Arg Gly Pro Ala
50 55 60
Ile Gly Gly Gln Trp Arg Gly Phe Trp Glu Ala Ser Ser Arg Gly
65 70 75
Gly Gly Ala Phe Ser Gly Gly Glu Asp Ala Ser Glu Gly Gly Ala
80 85 90
Glu Glu Gly Ala Gly Gly Ala Gly Gly Ser Ala Gly Ala Gly Glu
95 100 105
Gly Pro Val Ile Thr Ala Leu Thr Pro Met Thr Ile Pro Asp Val
110 115 120
Phe Pro His Leu Pro Leu Ile Ala Ile Thr Arg Asn Pro Val Phe
125 130 135
Pro Arg Phe Ile Lys Ile Ile Glu Val Lys Asn Lys Lys Leu Val
140 145 150

Glu Leu Leu Arg Arg Lys Val Arg Leu Ala Gln Pro Tyr Val Gly
155 160 165
Val Phe Leu Lys Arg Asp Asp Ser Asn Glu Ser Asp Val Val Glu
170 175 180
Ser Leu Asp Glu Ile Tyr His Pro Gly Thr Ala Gly Gly Gly Cys
185 190 195
Gln Ala Leu Ser Gly Pro Glu Leu Ser Ala Val Gly Ser Ala Pro
200 205 210
Gly Pro Gly Ser Gly Ala Thr Glu Arg Ala Ala Arg Ser Ser Asp
215 220 225
Pro Asp Pro Arg Asp Leu Ser Arg Leu Asn Gln Ser Val Ala
230 235

<210> 42
<211> 936
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509936CD1

<400> 42
Met Ala Thr Val Gly Glu Trp Ser Cys Val Arg Cys Thr Phe Leu
1 5 10 15
Asn Pro Ala Gly Gln Arg Gln Cys Ser Ile Cys Glu Ala Pro Arg
20 25 30
His Lys Pro Asp Leu Asn His Ile Leu Arg Leu Ser Val Glu Glu
35 40 45
Gln Lys Trp Pro Cys Ala Arg Cys Thr Phe Arg Asn Phe Leu Gly
50 55 60
Lys Glu Ala Cys Glu Val Cys Gly Phe Thr Pro Glu Pro Ala Pro
65 70 75
Gly Ala Ala Phe Leu Pro Val Leu Asn Gly Val Leu Pro Lys Pro
80 85 90
Pro Ala Ile Leu Gly Glu Pro Lys Gly Ser Cys Gln Glu Glu Ala
95 100 105
Gly Pro Val Arg Thr Ala Gly Leu Val Ala Thr Glu Pro Ala Arg
110 115 120
Gly Gln Cys Glu Asp Lys Asp Glu Glu Glu Lys Glu Glu Gln Glu
125 130 135
Glu Glu Glu Gly Ala Ala Glu Pro Arg Gly Gly Trp Ala Cys Pro
140 145 150
Arg Cys Thr Leu His Asn Thr Pro Val Ala Ser Ser Cys Ser Val
155 160 165
Cys Gly Gly Pro Arg Arg Leu Ser Leu Pro Arg Ile Pro Pro Glu
170 175 180
Ala Leu Val Val Pro Glu Val Val Ala Pro Ala Gly Phe His Val
185 190 195
Val Pro Ala Ala Pro Pro Pro Gly Leu Pro Gly Glu Gly Ala Glu
200 205 210
Ala Asn Pro Pro Ala Thr Ser Gln Gly Pro Ala Ala Glu Pro Glu
215 220 225
Pro Pro Arg Val Pro Pro Phe Ser Pro Phe Ser Ser Thr Leu Gln
230 235 240
Asn Asn Pro Val Pro Arg Ser Arg Arg Glu Val Pro Pro Gln Leu
245 250 255
Gln Pro Pro Val Pro Glu Ala Ala Gln Pro Ser Pro Ser Ala Gly
260 265 270
Cys Arg Gly Ala Pro Gln Gly Ser Gly Trp Ala Gly Ala Ser Arg
275 280 285
Leu Ala Glu Leu Leu Ser Gly Lys Arg Leu Ser Val Leu Glu Glu
290 295 300

Glu Ala Thr Glu Gly Gly Thr Ser Arg Val Glu Ala Gly Ser Ser
305 310 315
Thr Ser Gly Ser Asp Ile Ile Asp Leu Ala Gly Asp Thr Val Arg
320 325 330
Tyr Thr Pro Ala Ser Pro Ser Ser Pro Asp Phe Thr Thr Trp Ser
335 340 345
Cys Ala Lys Cys Thr Leu Arg Asn Pro Thr Val Ala Pro Arg Cys
350 355 360
Ser Ala Cys Gly Cys Ser Lys Leu His Gly Phe Gln Glu His Gly
365 370 375
Glu Pro Pro Thr His Cys Pro Asp Cys Gly Ala Asp Lys Pro Ser
380 385 390
Pro Cys Gly Arg Ser Cys Gly Arg Val Ser Ser Ala Gln Lys Ala
395 400 405
Ala Arg Val Leu Pro Glu Arg Pro Gly Gln Trp Ala Cys Pro Ala
410 415 420
Cys Thr Leu Leu Asn Ala Leu Arg Ala Lys His Cys Ala Ala Cys
425 430 435
His Thr Pro Gln Leu Leu Val Ala Gln Arg Arg Gly Ala Ala Pro
440 445 450
Leu Arg Arg Arg Glu Ser Met His Val Glu Gln Arg Arg Gln Thr
455 460 465
Asp Glu Gly Glu Ala Lys Ala Leu Trp Glu Asn Ile Val Ala Phe
470 475 480
Cys Arg Glu Asn Asn Val Ser Phe Val Asp Asp Ser Phe Pro Pro
485 490 495
Gly Pro Glu Ser Val Gly Phe Pro Ala Gly Asp Ser Val Gln Gln
500 505 510
Arg Val Arg Gln Trp Leu Arg Pro Gln Glu Ile Asn Cys Ser Val
515 520 525
Phe Arg Asp His Arg Ala Thr Trp Ser Val Phe His Thr Leu Arg
530 535 540
Pro Ser Asp Ile Leu Gln Gly Leu Leu Gly Asn Cys Trp Phe Leu
545 550 555
Ser Ala Leu Ala Val Leu Ala Glu Arg Pro Asp Leu Val Glu Arg
560 565 570
Val Met Val Thr Arg Ser Leu Cys Ala Glu Gly Ala Tyr Gln Val
575 580 585
Arg Leu Cys Lys Asp Gly Thr Trp Thr Thr Val Leu Val Asp Asp
590 595 600
Met Leu Pro Cys Asp Glu Ala Gly Cys Leu Leu Phe Ser Gln Ala
605 610 615
Gln Arg Lys Gln Leu Trp Val Ala Leu Ile Glu Lys Ala Leu Ala
620 625 630
Lys Leu His Gly Ser Tyr Phe Ala Leu Gln Ala Gly Arg Ala Ile
635 640 645
Glu Gly Leu Ala Thr Leu Thr Gly Ala Pro Cys Glu Ser Leu Ala
650 655 660
Leu Gln Leu Ser Ser Thr Asn Pro Arg Glu Glu Pro Val Asp Thr
665 670 675
Asp Leu Ile Trp Ala Lys Met Leu Ser Ser Lys Glu Ala Gly Phe
680 685 690
Leu Met Gly Ala Ser Cys Gly Gly Gly Asn Met Lys Val Asp Asp
695 700 705
Ser Ala Tyr Glu Ser Leu Gly Leu Arg Pro Arg His Ala Tyr Ser
710 715 720
Ile Leu Asp Val Arg Asp Val Gln Gly Thr Arg Leu Leu Arg Leu
725 730 735
Arg Asn Pro Trp Gly Arg Phe Ser Trp Asn Gly Ser Trp Ser Asp
740 745 750
Glu Trp Pro His Trp Pro Gly His Leu Arg Gly Glu Leu Met Pro
755 760 765
His Gly Ser Ser Glu Gly Val Phe Trp Met Glu Tyr Gly Asp Phe

```
                    770                         775                         780
Val Arg Tyr Phe Asp Ser Val Asp Ile Cys Lys Val His Ser Asp
                    785                         790                         795
Trp Gln Glu Ala Arg Val Gln Gly Cys Phe Pro Ser Ser Ala Ser
                    800                         805                         810
Ala Pro Val Gly Val Thr Ala Leu Thr Val Leu Glu Arg Ala Ser
                    815                         820                         825
Leu Glu Phe Ala Leu Phe Gln Glu Gly Ser Arg Arg Ser Asp Ala
                    830                         835                         840
Val Asp Ser His Leu Leu Asp Leu Cys Ile Leu Val Phe Arg Ala
                    845                         850                         855
Thr Phe Gly Ser Gly Gly His Leu Ser Leu Gly Arg Leu Leu Ala
                    860                         865                         870
His Ser Lys Arg Ala Val Lys Lys Phe Val Ser Cys Asp Val Met
                    875                         880                         885
Leu Glu Pro Gly Glu Tyr Ala Val Val Cys Cys Ala Phe Asn His
                    890                         895                         900
Trp Gly Pro Pro Leu Pro Gly Thr Pro Ala Pro Gln Gly Thr Trp
                    905                         910                         915
Pro Leu Pro Gln Ala His Ala Pro Gly Pro Arg Gly Gly Leu Arg
                    920                         925                         930
Phe Thr Arg Pro Ser Leu
                    935
```

<210> 43
<211> 1136
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509986CD1

<400> 43

```
Met Ala Ala Val Arg Gly Ala Pro Leu Leu Ser Cys Leu Leu Ala
  1                 5                  10                  15
Leu Leu Ala Leu Cys Pro Gly Gly Arg Pro Gln Thr Val Leu Thr
                   20                  25                  30
Asp Asp Glu Ile Glu Glu Phe Leu Glu Gly Phe Leu Ser Glu Leu
                   35                  40                  45
Glu Pro Glu Pro Arg Glu Asp Asp Val Glu Ala Pro Pro Pro Pro
                   50                  55                  60
Glu Pro Thr Pro Arg Val Arg Lys Ala Gln Ala Gly Gly Lys Pro
                   65                  70                  75
Gly Lys Arg Pro Gly Thr Ala Ala Glu Val Pro Pro Glu Lys Thr
                   80                  85                  90
Lys Asp Lys Gly Lys Lys Gly Lys Lys Asp Lys Gly Pro Lys Val
                   95                 100                 105
Pro Lys Glu Ser Leu Glu Gly Ser Pro Arg Pro Pro Lys Lys Gly
                  110                 115                 120
Lys Glu Lys Pro Pro Lys Ala Thr Lys Lys Pro Lys Glu Lys Pro
                  125                 130                 135
Pro Lys Ala Thr Lys Lys Pro Lys Glu Lys Pro Pro Lys Ala Thr
                  140                 145                 150
Lys Lys Pro Lys Glu Lys Pro Pro Lys Ala Thr Lys Lys Pro Pro
                  155                 160                 165
Ser Gly Lys Arg Pro Pro Ile Leu Ala Pro Ser Glu Thr Leu Glu
                  170                 175                 180
Trp Pro Leu Pro Pro Pro Pro Ser Pro Gly Pro Glu Glu Leu Pro
                  185                 190                 195
Gln Glu Gly Gly Ala Pro Leu Ser Asn Asn Trp Gln Asn Pro Gly
                  200                 205                 210
Glu Glu Thr His Val Glu Ala Arg Glu His Gln Pro Glu Pro Glu
```

215 220 225
Glu Glu Thr Glu Gln Pro Thr Leu Asp Tyr Asn Asp Gln Ile Glu
230 235 240
Arg Glu Asp Tyr Glu Asp Phe Glu Tyr Ile Arg Arg Gln Lys Gln
245 250 255
Pro Arg Pro Pro Pro Ser Arg Arg Arg Arg Pro Glu Arg Val Trp
260 265 270
Pro Glu Thr Pro Glu Glu Lys Ala Pro Ala Pro Ala Pro Glu Glu
275 280 285
Arg Ile Glu Pro Pro Val Lys Pro Leu Leu Pro Pro Leu Pro Pro
290 295 300
Asp Tyr Gly Asp Gly Tyr Val Ile Pro Asn Tyr Asp Asp Met Asp
305 310 315
Tyr Tyr Phe Gly Pro Pro Pro Pro Gln Lys Pro Asp Ala Glu Arg
320 325 330
Gln Thr Asp Glu Glu Lys Glu Glu Leu Lys Lys Pro Lys Lys Glu
335 340 345
Asp Ser Ser Pro Lys Glu Glu Thr Asp Lys Trp Ala Val Glu Lys
350 355 360
Gly Lys Asp His Lys Glu Pro Arg Lys Gly Glu Glu Leu Glu Glu
365 370 375
Glu Trp Thr Pro Thr Glu Lys Val Lys Cys Pro Pro Ile Gly Met
380 385 390
Glu Ser His Arg Ile Glu Asp Asn Gln Ile Arg Ala Ser Ser Met
395 400 405
Leu Arg His Gly Leu Gly Ala Gln Arg Gly Arg Leu Asn Met Gln
410 415 420
Thr Gly Ala Thr Glu Asp Asp Tyr Tyr Asp Gly Ala Trp Cys Ala
425 430 435
Glu Asp Asp Ala Arg Thr Gln Trp Ile Glu Val Asp Thr Arg Arg
440 445 450
Thr Thr Arg Phe Thr Gly Val Ile Thr Gln Gly Arg Asp Ser Ser
455 460 465
Ile His Asp Asp Phe Val Thr Thr Phe Phe Val Gly Phe Ser Asn
470 475 480
Asp Ser Gln Thr Trp Val Met Tyr Thr Asn Gly Tyr Glu Glu Met
485 490 495
Thr Phe His Gly Asn Val Asp Lys Asp Thr Pro Val Leu Ser Glu
500 505 510
Leu Pro Glu Pro Val Val Ala Arg Phe Ile Arg Ile Tyr Pro Leu
515 520 525
Thr Trp Asn Gly Ser Leu Cys Met Arg Leu Glu Val Leu Gly Cys
530 535 540
Ser Val Ala Pro Val Tyr Ser Tyr Tyr Ala Gln Asn Glu Val Val
545 550 555
Ala Thr Asp Asp Leu Asp Phe Arg His His Ser Tyr Lys Asp Met
560 565 570
Arg Gln Leu Met Lys Val Val Asn Glu Glu Cys Pro Thr Ile Thr
575 580 585
Arg Thr Tyr Ser Leu Gly Lys Ser Ser Arg Gly Leu Lys Ile Tyr
590 595 600
Ala Met Glu Ile Ser Asp Asn Pro Gly Glu His Glu Leu Gly Glu
605 610 615
Pro Glu Phe Arg Tyr Thr Ala Gly Ile His Gly Asn Glu Val Leu
620 625 630
Gly Arg Glu Leu Leu Leu Leu Leu Met Gln Tyr Leu Cys Arg Glu
635 640 645
Tyr Arg Asp Gly Asn Pro Arg Val Arg Ser Leu Val Gln Asp Thr
650 655 660
Arg Ile His Leu Val Pro Ser Leu Asn Pro Asp Gly Tyr Glu Val
665 670 675
Ala Ala Gln Met Gly Ser Glu Phe Gly Asn Trp Ala Leu Gly Leu
680 685 690

Trp Thr Glu Glu Gly Phe Asp Ile Phe Glu Asp Phe Pro Asp Leu
695 700 705
Asn Ser Val Leu Trp Gly Ala Glu Glu Arg Lys Trp Val Pro Tyr
710 715 720
Arg Val Pro Asn Asn Asn Leu Pro Ile Pro Glu Arg Tyr Leu Ser
725 730 735
Pro Asp Ala Thr Val Ser Thr Glu Val Arg Ala Ile Ile Ala Trp
740 745 750
Met Glu Lys Asn Pro Phe Val Leu Gly Ala Asn Leu Asn Gly Gly
755 760 765
Glu Arg Leu Val Ser Tyr Pro Tyr Asp Met Ala Arg Thr Pro Thr
770 775 780
Gln Glu Gln Leu Leu Ala Ala Ala Met Ala Ala Ala Arg Gly Glu
785 790 795
Asp Glu Asp Glu Val Ser Glu Ala Gln Glu Thr Pro Asp His Ala
800 805 810
Ile Phe Arg Trp Leu Ala Ile Ser Phe Ala Ser Ala His Leu Thr
815 820 825
Leu Thr Glu Pro Tyr Arg Gly Gly Cys Gln Ala Gln Asp Tyr Thr
830 835 840
Gly Gly Met Gly Ile Val Asn Gly Ala Lys Trp Asn Pro Arg Thr
845 850 855
Gly Thr Ile Asn Asp Phe Ser Tyr Leu His Thr Asn Cys Leu Glu
860 865 870
Leu Ser Phe Tyr Leu Gly Cys Asp Lys Phe Pro His Glu Ser Glu
875 880 885
Leu Pro Arg Glu Trp Glu Asn Asn Lys Glu Ala Leu Leu Thr Phe
890 895 900
Met Glu Gln Val His Arg Gly Ile Asn His Gly Val Lys Thr Ala
905 910 915
Ser Gly Gly Asp Tyr Trp Arg Ile Leu Asn Pro Gly Glu Tyr Arg
920 925 930
Val Thr Ala His Ala Glu Gly Tyr Thr Pro Ser Ala Lys Thr Cys
935 940 945
Asn Val Asp Tyr Asp Ile Gly Ala Thr Gln Cys Asn Phe Ile Leu
950 955 960
Ala Arg Ser Asn Trp Lys Arg Ile Arg Glu Ile Met Ala Met Asn
965 970 975
Gly Asn Arg Pro Ile Pro His Ile Asp Pro Ser Arg Pro Met Thr
980 985 990
Pro Gln Gln Arg Arg Leu Gln Gln Arg Arg Leu Gln His Arg Leu
995 1000 1005
Arg Leu Arg Ala Gln Met Arg Leu Arg Arg Leu Asn Ala Thr Thr
1010 1015 1020
Thr Leu Gly Pro His Thr Val Pro Pro Thr Leu Pro Pro Ala Pro
1025 1030 1035
Ala Thr Thr Leu Ser Thr Thr Ile Glu Pro Trp Gly Leu Ile Pro
1040 1045 1050
Pro Thr Thr Ala Gly Trp Glu Glu Ser Glu Thr Glu Thr Tyr Thr
1055 1060 1065
Glu Val Val Thr Glu Phe Gly Thr Glu Val Glu Pro Glu Phe Gly
1070 1075 1080
Thr Lys Val Glu Pro Glu Phe Glu Thr Gln Leu Glu Pro Glu Phe
1085 1090 1095
Glu Thr Gln Leu Glu Pro Glu Phe Glu Glu Glu Glu Glu Glu Glu
1100 1105 1110
Lys Glu Glu Glu Ile Ala Thr Gly Gln Ala Phe Pro Phe Thr Thr
1115 1120 1125
Val Glu Thr Tyr Thr Val Asn Phe Gly Asp Phe
1130 1135

<210> 44
<211> 617

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510010CD1

<400> 44
Met Ala Arg Gly Tyr Gly Ala Thr Val Ser Leu Val Leu Leu Gly
1 5 10 15
Leu Gly Leu Ala Leu Ala Val Ile Val Leu Ala Val Val Leu Ser
20 25 30
Arg His Gln Ala Pro Cys Gly Pro Gln Ala Phe Ala His Ala Ala
35 40 45
Val Ala Ala Asp Ser Lys Val Cys Ser Asp Ile Gly Arg Ala Ile
50 55 60
Leu Gln Gln Gln Gly Ser Pro Val Asp Ala Thr Ile Ala Ala Leu
65 70 75
Val Cys Thr Ser Val Val Asn Pro Gln Ser Met Gly Leu Gly Gly
80 85 90
Gly Val Ile Phe Thr Ile Tyr Asn Val Thr Thr Gly Lys Val Glu
95 100 105
Val Ile Asn Ala Arg Glu Thr Val Pro Ala Ser His Ala Pro Ser
110 115 120
Leu Leu Asp Gln Cys Ala Gln Ala Leu Pro Leu Gly Thr Gly Ala
125 130 135
Gln Trp Ile Gly Val Pro Gly Glu Leu Arg Gly Tyr Ala Glu Ala
140 145 150
His Arg Arg His Gly Arg Leu Pro Trp Ala Gln Leu Phe Gln Pro
155 160 165
Thr Ile Ala Leu Leu Arg Gly Gly His Val Val Ala Pro Val Leu
170 175 180
Ser Arg Phe Leu His Asn Ser Ile Leu Arg Pro Ser Leu Gln Ala
185 190 195
Ser Thr Leu Arg Gln Leu Phe Phe Asn Gly Thr Glu Pro Leu Arg
200 205 210
Pro Gln Asp Pro Leu Pro Trp Pro Ala Leu Ala Thr Thr Leu Glu
215 220 225
Thr Val Ala Thr Glu Gly Val Glu Val Phe Tyr Thr Gly Arg Leu
230 235 240
Gly Gln Met Leu Val Glu Asp Ile Ala Lys Glu Gly Ser Gln Leu
245 250 255
Thr Leu Gln Asp Leu Ala Lys Phe Gln Pro Glu Val Val Asp Ala
260 265 270
Leu Glu Val Pro Leu Gly Asp Tyr Thr Leu Tyr Ser Pro Pro Pro
275 280 285
Pro Ala Gly Gly Ala Ile Leu Ser Phe Ile Leu Asn Val Leu Arg
290 295 300
Gly Phe Asn Phe Ser Thr Glu Ser Met Ala Arg Pro Glu Gly Arg
305 310 315
Val Asn Val Tyr His His Leu Val Glu Thr Leu Lys Phe Ala Arg
320 325 330
Gly Gln Arg Trp Arg Leu Gly Asp Pro Arg Ser His Pro Lys Leu
335 340 345
Gln Asn Ala Ser Arg Asp Leu Leu Gly Glu Thr Leu Ala Gln Leu
350 355 360
Ile Arg Gln Gln Ile Asp Gly Arg Gly Asp His Gln Leu Ser His
365 370 375
Tyr Ser Leu Ala Glu Ala Trp Gly His Gly Thr Gly Thr Ser His
380 385 390
Val Ser Val Leu Gly Glu Asp Gly Ser Ala Val Ala Ala Thr Ser
395 400 405
Thr Ile Asn Thr Pro Phe Gly Ala Met Val Tyr Ser Pro Arg Thr

```
                410                 415                 420
Gly Ile Ile Leu Asn Asn Glu Leu Leu Asp Leu Cys Glu Arg Cys
                425                 430                 435
Pro Arg Gly Ser Gly Thr Thr Pro Ser Pro Val Ser Gly Asp Arg
                440                 445                 450
Val Gly Gly Ala Pro Gly Arg Cys Trp Pro Pro Val Pro Gly Glu
                455                 460                 465
Arg Ser Pro Ser Ser Met Val Pro Ser Ile Leu Ile Asn Lys Ala
                470                 475                 480
Gln Gly Ser Lys Leu Val Ile Gly Gly Ala Gly Gly Glu Leu Ile
                485                 490                 495
Ile Ser Ala Val Ala Gln Ala Ile Met Ser Lys Leu Trp Leu Gly
                500                 505                 510
Phe Asp Leu Arg Ala Ala Ile Ala Ala Pro Ile Leu His Val Asn
                515                 520                 525
Ser Lys Gly Cys Val Glu Tyr Glu Pro Asn Phe Ser Gln Lys Gln
                530                 535                 540
Asp Leu Ala Ala Leu Pro Arg Leu Val Ser Asn Ser Trp Pro Gln
                545                 550                 555
Val Ile Leu Leu Pro Gln Pro Pro Lys Leu Leu Ala Leu Gln Glu
                560                 565                 570
Val Gln Arg Gly Leu Gln Asp Arg Gly Gln Asn Gln Thr Gln Arg
                575                 580                 585
Pro Phe Phe Leu Asn Val Val Gln Ala Val Ser Gln Glu Gly Ala
                590                 595                 600
Cys Val Tyr Ala Val Ser Asp Leu Arg Lys Ser Gly Glu Ala Ala
                605                 610                 615
Gly Tyr


<210> 45
<211> 316
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510056CD1

<400> 45
Met Cys Gln Trp Ala His Pro Ser Leu Gly Pro His Ser Glu Pro
 1               5                  10                  15
Ala Ser Pro Ser Val Glu Tyr Ile Arg Arg Gln Lys Gln Pro Arg
                20                  25                  30
Pro Pro Pro Ser Arg Arg Arg Arg Pro Glu Arg Val Trp Pro Asp
                35                  40                  45
Pro Pro Glu Glu Lys Ala Pro Ala Pro Ala Pro Glu Glu Arg Ile
                50                  55                  60
Glu Pro Pro Val Lys Pro Leu Leu Pro Pro Leu Pro Pro Asp Tyr
                65                  70                  75
Gly Asp Gly Tyr Val Ile Pro Asn Tyr Asp Asp Met Asp Tyr Tyr
                80                  85                  90
Phe Gly Pro Pro Pro Pro Gln Lys Pro Asp Ala Glu Arg Gln Thr
                95                 100                 105
Asp Glu Glu Lys Glu Glu Leu Lys Lys Pro Lys Lys Glu Asp Ser
               110                 115                 120
Ser Pro Lys Glu Glu Thr Asp Lys Trp Ala Val Glu Lys Gly Lys
               125                 130                 135
Asp His Lys Glu Pro Arg Lys Gly Glu Glu Leu Glu Glu Glu Trp
               140                 145                 150
Thr Pro Thr Glu Lys Val Lys Cys Pro Pro Ile Gly Met Glu Ser
               155                 160                 165
His Arg Ile Glu Asp Asn Gln Ile Arg Ala Ser Ser Met Leu Arg
```

170 175 180
His Gly Leu Gly Ala Gln Arg Gly Arg Leu Asn Met Gln Thr Gly
185 190 195
Ala Thr Glu Asp Asp Tyr Tyr Asp Gly Ala Trp Cys Ala Glu Asp
200 205 210
Asp Ala Arg Thr Gln Trp Ile Glu Val Asp Thr Arg Arg Thr Thr
215 220 225
Arg Phe Thr Gly Val Ile Thr Gln Gly Arg Asp Ser Ser Ile His
230 235 240
Asp Asp Phe Val Thr Thr Phe Phe Val Gly Phe Ser Asn Asp Ser
245 250 255
Gln Thr Trp Val Met Tyr Thr Asn Gly Tyr Glu Glu Met Val Gly
260 265 270
Thr Met Pro Arg Leu Leu Ala Leu Leu Pro Leu Ser Gly Arg Gly
275 280 285
Val Gly Ser Gln Arg Gly Trp Gln Tyr Cys Ser Glu Ala Cys Leu
290 295 300
Ser Pro Asp Leu Ser Trp Glu Arg Gly Gln Gly His Thr Arg Ala
305 310 315
Glu

<210> 46
<211> 418
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510398CD1

<400> 46
Met Glu Arg Asp Ser His Gly Asn Ala Ser Pro Ala Arg Thr Pro
1 5 10 15
Ser Ala Gly Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr Pro
20 25 30
Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser
35 40 45
Pro Ala Gly Thr Pro Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser
50 55 60
Pro Ala Gly Thr Pro Pro Gly Arg Ala Ser Pro Gly Arg Ala Ser
65 70 75
Pro Ala Gln Ala Ser Pro Ala Arg Ala Ser Pro Ala Leu Ala Ser
80 85 90
Leu Ser Arg Ser Ser Ser Gly Arg Ser Ser Ser Ala Arg Ser Ala
95 100 105
Ser Val Thr Thr Ser Pro Thr Arg Val Tyr Leu Val Arg Ala Thr
110 115 120
Pro Val Gly Ala Val Pro Ile Arg Ser Ser Pro Ala Arg Ser Ala
125 130 135
Pro Ala Thr Arg Ala Thr Arg Glu Ser Pro Gly Thr Ser Leu Pro
140 145 150
Lys Phe Thr Trp Arg Glu Gly Gln Lys Gln Leu Pro Leu Ile Gly
155 160 165
Cys Val Leu Leu Leu Ile Ala Leu Val Val Ser Leu Ile Ile Leu
170 175 180
Phe Gln Phe Trp Gln Gly His Thr Gly Ile Arg Tyr Lys Glu Gln
185 190 195
Arg Glu Ser Cys Pro Lys His Ala Val Arg Cys Asp Gly Val Val
200 205 210
Asp Cys Lys Leu Lys Ser Asp Glu Leu Gly Cys Val Arg Phe Asp
215 220 225
Trp Asp Lys Ser Leu Leu Lys Ile Tyr Ser Gly Ser Ser His Gln

230 235 240
Trp Leu Pro Ile Cys Ser Ser Asn Trp Asn Asp Ser Tyr Ser Glu
245 250 255
Lys Thr Cys Gln Gln Leu Gly Phe Glu Ser Ala His Arg Thr Thr
260 265 270
Glu Val Ala His Arg Asp Phe Ala Asn Ser Phe Ser Ile Leu Arg
275 280 285
Tyr Asn Ser Thr Ile Gln Glu Ser Leu His Arg Ser Glu Cys Pro
290 295 300
Ser Gln Arg Tyr Ile Ser Leu Gln Cys Ser His Cys Gly Leu Arg
305 310 315
Ala Met Thr Gly Arg Ile Val Gly Gly Ala Leu Ala Ser Asp Ser
320 325 330
Lys Trp Pro Trp Gln Val Ser Leu His Phe Gly Thr Thr His Ile
335 340 345
Cys Gly Gly Thr Leu Ile Asp Ala Gln Trp Val Leu Thr Ala Ala
350 355 360
His Cys Phe Phe Val Thr Arg Glu Lys Val Leu Glu Gly Trp Lys
365 370 375
Val Tyr Ala Gly Thr Ser Asn Ser Tyr Pro Gly Pro Lys Ala Ser
380 385 390
Ala Gly Gln Lys Ser Lys Thr Leu Lys Asp Pro Tyr Met Glu His
395 400 405
Phe Cys Phe Ile Ile Arg Glu Thr Glu Ala Gln Gly Leu
410 415

<210> 47
<211> 543
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510498CD1

<400> 47
Met Ala Ala Val Arg Gly Ala Pro Leu Leu Ser Cys Leu Leu Ala
1 5 10 15
Leu Leu Ala Leu Cys Pro Gly Gly Arg Pro Gln Thr Val Leu Thr
20 25 30
Asp Asp Glu Ile Glu Glu Phe Leu Glu Gly Phe Leu Ser Glu Leu
35 40 45
Glu Pro Glu Pro Arg Glu Asp Asp Val Glu Ala Pro Pro Pro Pro
50 55 60
Glu Pro Thr Pro Arg Val Arg Lys Ala Gln Ala Gly Gly Lys Pro
65 70 75
Gly Lys Arg Pro Gly Thr Ala Ala Glu Val Pro Pro Glu Lys Thr
80 85 90
Lys Asp Lys Gly Lys Lys Gly Lys Lys Asp Lys Gly Pro Lys Val
95 100 105
Pro Lys Glu Ser Leu Glu Gly Ser Pro Arg Pro Pro Lys Lys Gly
110 115 120
Lys Glu Lys Pro Pro Lys Ala Thr Lys Lys Pro Lys Glu Lys Pro
125 130 135
Pro Lys Ala Thr Lys Lys Pro Lys Glu Lys Pro Pro Lys Ala Thr
140 145 150
Lys Lys Pro Lys Glu Lys Pro Pro Lys Ala Thr Lys Lys Pro Pro
155 160 165
Ser Gly Lys Arg Pro Pro Ile Leu Ala Pro Ser Glu Thr Leu Glu
170 175 180
Trp Pro Leu Pro Pro Pro Pro Ser Pro Gly Pro Glu Glu Leu Pro
185 190 195
Gln Glu Gly Gly Ala Pro Leu Ser Asn Asn Trp Gln Asn Pro Gly

200 205 210
Glu Glu Thr His Val Glu Ala Arg Glu His Arg Pro Glu Pro Glu
215 220 225
Glu Glu Thr Glu Gln Pro Thr Leu Asp Tyr Asn Asp Gln Ile Glu
230 235 240
Arg Glu Asp Tyr Glu Asp Phe Glu Tyr Ile Arg Arg Gln Lys Gln
245 250 255
Pro Arg Pro Pro Pro Ser Arg Arg Arg Arg Pro Glu Arg Val Trp
260 265 270
Pro Glu Thr Pro Glu Glu Lys Ala Pro Ala Pro Ala Pro Glu Glu
275 280 285
Arg Ile Glu Pro Pro Val Lys Pro Leu Leu Pro Pro Leu Pro Pro
290 295 300
Asp Tyr Gly Asp Gly Tyr Val Ile Pro Asn Tyr Asp Asp Met Asp
305 310 315
Tyr Tyr Phe Gly Pro Pro Pro Pro Gln Lys Pro Asp Ala Glu Arg
320 325 330
Gln Thr Asp Glu Glu Lys Glu Glu Leu Lys Lys Pro Lys Lys Glu
335 340 345
Asp Ser Ser Pro Lys Glu Glu Thr Asp Lys Trp Ala Val Glu Lys
350 355 360
Gly Lys Asp His Lys Glu Pro Arg Lys Gly Glu Glu Leu Glu Glu
365 370 375
Glu Trp Thr Pro Thr Glu Lys Val Lys Cys Pro Pro Ile Gly Met
380 385 390
Glu Ser His Arg Ile Glu Asp Asn Gln Ile Arg Ala Ser Ser Met
395 400 405
Leu Arg His Gly Leu Gly Ala Gln Arg Gly Arg Leu Asn Met Gln
410 415 420
Thr Gly Ala Thr Glu Asp Asp Tyr Tyr Asp Gly Ala Trp Cys Ala
425 430 435
Glu Asp Asp Ala Arg Thr Gln Trp Ile Glu Val Asp Thr Arg Arg
440 445 450
Thr Thr Arg Phe Thr Gly Val Ile Thr Gln Gly Arg Asp Ser Ser
455 460 465
Ile His Asp Asp Phe Val Thr Thr Phe Phe Val Gly Phe Ser Asn
470 475 480
Asp Ser Gln Thr Trp Val Met Tyr Thr Asn Gly Tyr Glu Glu Met
485 490 495
Val Gly Thr Met Pro Arg Leu Leu Ala Leu Leu Pro Leu Ser Gly
500 505 510
Arg Gly Val Gly Ser Gln Arg Gly Trp Gln Tyr Cys Ser Glu Ala
515 520 525
Cys Leu Ser Pro Asp Leu Ser Trp Glu Arg Gly Gln Gly His Thr
530 535 540
Arg Ala Glu

<210> 48
<211> 742
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510044CD1

<400> 48
Met Ala Thr Ala Gly Gly Gly Ser Gly Ala Asp Pro Gly Ser Arg
1 5 10 15
Gly Leu Leu Arg Leu Leu Ser Phe Cys Val Leu Leu Ala Gly Leu
20 25 30
Phe Leu Asp Phe Glu Pro Asp Pro Leu Pro Leu Cys Pro Asp Val

35 40 45
Ser Phe Phe Pro Gln Pro Gln Pro Pro Thr Pro Gln Ser Lys Gly
50 55 60
Thr Gly Leu Cys Arg Gly Asn Ser Val Glu Arg Lys Ile Tyr Ile
65 70 75
Pro Leu Asn Lys Thr Ala Pro Cys Val Arg Leu Leu Asn Ala Thr
80 85 90
His Gln Ile Gly Cys Gln Ser Ser Ile Ser Gly Asp Thr Gly Val
95 100 105
Ile His Val Val Glu Lys Glu Glu Asp Leu Gln Trp Val Leu Thr
110 115 120
Asp Gly Pro Asn Pro Pro Tyr Met Val Leu Leu Glu Ser Lys His
125 130 135
Phe Thr Arg Asp Leu Met Glu Lys Leu Lys Gly Arg Thr Ser Arg
140 145 150
Ile Ala Gly Leu Ala Val Ser Leu Thr Lys Pro Ser Pro Ala Ser
155 160 165
Gly Phe Ser Pro Ser Val Gln Cys Pro Asn Asp Gly Phe Gly Val
170 175 180
Tyr Ser Asn Ser Tyr Gly Pro Glu Phe Ala His Cys Arg Glu Ile
185 190 195
Gln Trp Asn Ser Leu Gly Asn Gly Leu Ala Tyr Glu Asp Phe Ser
200 205 210
Phe Pro Ile Phe Leu Leu Glu Asp Glu Asn Glu Thr Lys Val Ile
215 220 225
Lys Gln Cys Tyr Gln Asp His Asn Leu Ser Gln Asn Gly Ser Ala
230 235 240
Pro Thr Phe Pro Leu Cys Ala Met Gln Leu Phe Ser His Met His
245 250 255
Ala Val Ile Ser Thr Ala Thr Cys Met Arg Arg Arg Ser Ile Gln
260 265 270
Ser Thr Phe Ser Ile Asn Pro Glu Ile Val Cys Asp Pro Leu Ser
275 280 285
Asp Tyr Asn Val Trp Ser Met Leu Lys Pro Ile Asn Thr Thr Gly
290 295 300
Thr Leu Lys Pro Asp Asp Arg Val Val Val Ala Ala Thr Arg Leu
305 310 315
Asp Ser Arg Ser Phe Phe Trp Asn Val Ala Pro Gly Ala Glu Ser
320 325 330
Ala Val Ala Ser Phe Val Thr Gln Leu Ala Ala Ala Glu Ala Leu
335 340 345
Gln Lys Ala Pro Asp Val Thr Thr Leu Pro Arg Asn Val Met Phe
350 355 360
Val Phe Phe Gln Gly Glu Thr Phe Asp Tyr Ile Gly Ser Ser Arg
365 370 375
Met Val Tyr Asp Met Glu Lys Gly Lys Phe Pro Val Gln Leu Glu
380 385 390
Asn Val Asp Ser Phe Val Glu Leu Gly Gln Val Ala Leu Arg Thr
395 400 405
Ser Leu Glu Leu Trp Met His Thr Asp Pro Val Ser Gln Lys Asn
410 415 420
Glu Ser Val Arg Asn Gln Val Glu Asp Leu Leu Ala Thr Leu Glu
425 430 435
Lys Ser Gly Ala Gly Val Pro Ala Val Ile Leu Arg Arg Pro Asn
440 445 450
Gln Ser Gln Pro Leu Pro Pro Ser Ser Leu Gln Arg Phe Leu Arg
455 460 465
Ala Arg Asn Ile Ser Gly Val Val Leu Ala Asp His Ser Gly Ala
470 475 480
Phe His Asn Lys Tyr Tyr Gln Ser Ile Tyr Asp Thr Ala Glu Asn
485 490 495
Ile Asn Val Ser Tyr Pro Glu Trp Leu Ser Pro Glu Glu Asp Leu
500 505 510

Asn Phe Val Thr Asp Thr Ala Lys Ala Leu Ala Asp Val Ala Thr
515 520 525
Val Leu Gly Arg Ala Leu Tyr Glu Leu Ala Gly Gly Thr Asn Phe
530 535 540
Ser Asp Thr Val Gln Ala Asp Pro Gln Thr Val Thr Arg Leu Leu
545 550 555
Tyr Gly Phe Leu Ile Lys Ala Asn Asn Ser Trp Phe Gln Ser Ile
560 565 570
Leu Arg Gln Asp Leu Arg Ser Tyr Leu Gly Asp Gly Pro Leu Gln
575 580 585
His Tyr Ile Ala Val Ser Ser Pro Thr Asn Thr Thr Tyr Val Val
590 595 600
Gln Tyr Ala Leu Ala Asn Leu Thr Gly Thr Val Val Asn Leu Thr
605 610 615
Arg Glu Gln Cys Gln Asp Pro Ser Lys Val Pro Ser Glu Asn Lys
620 625 630
Asp Leu Tyr Glu Tyr Ser Trp Val Gln Gly Pro Leu His Ser Asn
635 640 645
Glu Thr Asp Arg Leu Pro Arg Cys Val Arg Ser Thr Ala Arg Leu
650 655 660
Ala Arg Ala Leu Ser Pro Ala Phe Glu Leu Ser Gln Trp Ser Ser
665 670 675
Thr Glu Tyr Ser Thr Trp Thr Glu Ser Arg Trp Lys Asp Ile Arg
680 685 690
Ala Arg Ile Phe Leu Ile Ala Ser Lys Glu Leu Glu Leu Ile Thr
695 700 705
Leu Thr Val Gly Phe Gly Ile Leu Ile Phe Ser Leu Ile Val Thr
710 715 720
Tyr Cys Ile Asn Ala Lys Ala Asp Val Leu Phe Ile Ala Pro Arg
725 730 735
Glu Pro Gly Ala Val Ser Tyr
740

<210> 49
<211> 47
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504509CD1

<400> 49
Met Ser Gly His Ser Pro Thr Arg Gly Ala Met Gln Val Ala Met
1 5 10 15
Asn Gly Lys Ala Arg Lys Glu Ala Val Gln Thr Ala Ala Lys Glu
20 25 30
Leu Leu Lys Phe Val Asn Arg Ser Pro Ser Pro Phe His Gly Trp
35 40 45
Leu Gln

<210> 50
<211> 74
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506825CD1

<400> 50
Met Thr Ser Lys Gly Pro Glu Glu Glu His Pro Ser Val Thr Leu

```
 1               5                   10                  15
Phe Arg Gln Tyr Leu Arg Ile Arg Thr Val Gln Pro Lys Pro Asp
                20                  25                  30
Tyr Gly Gly Thr Trp Leu Cys Gly Asp Arg Val Asp Leu Ala Arg
                35                  40                  45
His Gln Pro Tyr Thr Leu Leu His Leu Ala Gln Leu Pro His Gly
                50                  55                  60
Cys Gly Ala Cys Leu Gln Gly Thr Leu Glu Ser Arg Pro Leu
                65                  70

<210> 51
<211> 343
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506828CD1

<400> 51
Met Thr Ser Lys Gly Pro Glu Glu Glu His Pro Ser Val Thr Leu
 1               5                   10                  15
Phe Arg Gln Tyr Leu Arg Ile Arg Thr Val Gln Pro Lys Pro Asp
                20                  25                  30
Tyr Gly Ala Ala Val Ala Phe Phe Glu Glu Thr Ala Arg Gln Leu
                35                  40                  45
Gly Leu Gly Cys Gln Lys Val Glu Val Ala Pro Gly Tyr Val Val
                50                  55                  60
Thr Val Leu Thr Trp Pro Gly Thr Asn Pro Thr Leu Ser Ser Ile
                65                  70                  75
Leu Leu Asn Ser His Thr Asp Val Val Pro Val Phe Lys Glu His
                80                  85                  90
Trp Ser His Asp Pro Phe Glu Ala Phe Lys Asp Ser Glu Gly Tyr
                95                  100                 105
Ile Tyr Ala Arg Gly Ala Gln Asp Met Lys Cys Val Ser Ile Gln
                110                 115                 120
Tyr Leu Glu Ala Val Arg Arg Leu Lys Val Glu Gly His Arg Phe
                125                 130                 135
Pro Arg Thr Ile His Met Thr Phe Val Pro Asp Glu Glu Val Gly
                140                 145                 150
Gly His Gln Gly Met Glu Leu Phe Val Gln Arg Pro Glu Phe His
                155                 160                 165
Ala Leu Arg Ala Gly Phe Ala Leu Asp Glu Gly Ile Ala Asn Pro
                170                 175                 180
Thr Asp Ala Phe Thr Val Phe Tyr Ser Glu Arg Ser Pro Trp Trp
                185                 190                 195
Val Arg Val Thr Ser Thr Gly Arg Pro Gly His Ala Ser Arg Phe
                200                 205                 210
Met Glu Asp Thr Ala Ala Glu Lys Leu Ala Phe Glu Glu Gln Leu
                215                 220                 225
Gln Ser Trp Cys Gln Ala Ala Gly Glu Gly Val Thr Leu Glu Phe
                230                 235                 240
Ala Gln Lys Trp Met His Pro Gln Val Thr Pro Thr Asp Asp Ser
                245                 250                 255
Asn Pro Trp Trp Ala Ala Phe Ser Arg Val Cys Lys Asp Met Asn
                260                 265                 270
Leu Thr Leu Glu Pro Glu Ile Met Pro Ala Ala Thr Asp Asn Arg
                275                 280                 285
Tyr Ile Arg Ala Val Gly Val Pro Ala Leu Gly Phe Ser Pro Met
                290                 295                 300
Asn Arg Thr Pro Val Leu Leu His Asp His Asp Glu Arg Leu His
                305                 310                 315
Glu Ala Val Phe Leu Arg Gly Val Asp Ile Tyr Thr Arg Leu Leu
```

```
                320                 325                 330
Pro Ala Leu Ala Ser Val Pro Ala Leu Pro Ser Asp Ser
                335                 340

<210> 52
<211> 373
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506831CD1

<400> 52
Met Thr Ser Lys Gly Pro Glu Glu Glu His Pro Ser Val Thr Leu
 1               5                  10                  15
Phe Arg Gln Tyr Leu Arg Ile Arg Thr Val Gln Pro Lys Pro Asp
                20                  25                  30
Tyr Gly Thr Asn Pro Thr Leu Ser Ser Ile Leu Leu Asn Ser His
                35                  40                  45
Thr Asp Val Val Pro Val Phe Lys Glu His Trp Ser His Asp Pro
                50                  55                  60
Phe Glu Ala Phe Lys Asp Ser Glu Gly Tyr Ile Tyr Ala Arg Gly
                65                  70                  75
Ala Gln Asp Met Lys Cys Val Ser Ile Gln Tyr Leu Glu Ala Val
                80                  85                  90
Arg Arg Leu Lys Val Glu Gly His Arg Phe Pro Arg Thr Ile His
                95                 100                 105
Met Thr Phe Val Pro Asp Glu Glu Val Gly Gly His Gln Gly Met
               110                 115                 120
Glu Leu Phe Val Gln Arg Pro Glu Phe His Ala Leu Arg Ala Gly
               125                 130                 135
Phe Ala Leu Asp Glu Gly Ile Ala Asn Pro Thr Asp Ala Phe Thr
               140                 145                 150
Val Phe Tyr Ser Glu Arg Ser Pro Trp Trp Val Arg Val Thr Ser
               155                 160                 165
Thr Gly Arg Pro Gly His Ala Ser Arg Phe Met Glu Asp Thr Ala
               170                 175                 180
Ala Glu Lys Leu His Lys Val Val Asn Ser Ile Leu Ala Phe Arg
               185                 190                 195
Glu Lys Glu Trp Gln Arg Leu Gln Ser Asn Pro His Leu Lys Glu
               200                 205                 210
Gly Ser Val Thr Ser Val Asn Leu Thr Lys Leu Glu Gly Gly Val
               215                 220                 225
Ala Tyr Asn Val Ile Pro Ala Thr Met Ser Ala Ser Phe Asp Phe
               230                 235                 240
Arg Val Ala Pro Asp Val Asp Phe Lys Ala Phe Glu Glu Gln Leu
               245                 250                 255
Gln Ser Trp Cys Gln Ala Ala Gly Glu Gly Val Thr Leu Glu Phe
               260                 265                 270
Ala Gln Lys Trp Met His Pro Gln Val Thr Pro Thr Asp Asp Ser
               275                 280                 285
Asn Pro Trp Trp Ala Ala Phe Ser Arg Val Cys Lys Asp Met Asn
               290                 295                 300
Leu Thr Leu Glu Pro Glu Ile Met Pro Ala Ala Thr Asp Asn Arg
               305                 310                 315
Tyr Ile Arg Ala Val Gly Val Pro Ala Leu Gly Phe Ser Pro Met
               320                 325                 330
Asn Arg Thr Pro Val Leu Leu His Asp His Asp Glu Arg Leu His
               335                 340                 345
Glu Ala Val Phe Leu Arg Gly Val Asp Ile Tyr Thr Arg Leu Leu
               350                 355                 360
Pro Ala Leu Ala Ser Val Pro Ala Leu Pro Ser Asp Ser
```

365 370

<210> 53
<211> 180
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509968CD1

<400> 53
Met Trp Pro Leu Ala Leu Val Ile Ala Ser Leu Thr Leu Ala Leu
1 5 10 15
Ser Gly Gly Val Ser Gln Glu Ser Ser Lys Val Leu Asn Thr Asn
20 25 30
Gly Thr Ser Gly Phe Leu Pro Gly Gly Tyr Thr Cys Phe Pro His
35 40 45
Ser Gln Pro Trp Gln Ala Ala Leu Leu Val Gln Gly Arg Leu Leu
50 55 60
Cys Gly Gly Val Leu Val His Pro Lys Trp Val Leu Thr Ala Ala
65 70 75
His Cys Leu Lys Glu Gly Leu Lys Val Tyr Leu Gly Lys His Ala
80 85 90
Leu Gly Arg Val Glu Ala Gly Glu Gln Val Arg Glu Val Val His
95 100 105
Ser Ile Pro His Pro Glu Tyr Arg Arg Ser Pro Thr His Leu Asn
110 115 120
His Asp His Asp Ile Met Leu Leu Glu Leu Gln Ser Pro Val Gln
125 130 135
Leu Thr Gly Tyr Ile Gln Thr Leu Pro Leu Ser His Asn Asn Arg
140 145 150
Leu Thr Pro Gly Thr Thr Cys Arg Val Ser Gly Trp Gly Thr Thr
155 160 165
Thr Ser Pro Gln Gly Met His Pro His Arg Trp Pro Glu Ala Pro
170 175 180

<210> 54
<211> 293
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510232CD1

<400> 54
Met Asp Phe Ser Arg Asn Leu Tyr Asp Ile Gly Glu Gln Leu Asp
1 5 10 15
Ser Glu Asp Leu Ala Ser Leu Lys Phe Leu Ser Leu Asp Tyr Ile
20 25 30
Pro Gln Arg Lys Gln Glu Pro Ile Lys Asp Ala Leu Met Leu Phe
35 40 45
Gln Arg Leu Gln Glu Lys Arg Met Leu Glu Glu Ser Asn Leu Ser
50 55 60
Phe Leu Lys Glu Leu Leu Phe Arg Ile Asn Arg Leu Asp Leu Leu
65 70 75
Ile Thr Tyr Leu Asn Thr Arg Lys Glu Glu Met Glu Arg Glu Leu
80 85 90
Gln Thr Pro Gly Arg Ala Gln Ile Ser Ala Tyr Arg Phe His Phe
95 100 105
Cys Arg Met Ser Trp Ala Glu Ala Asn Ser Gln Cys Gln Thr Gln

```
                110                     115                     120
Ser Val Pro Phe Trp Arg Arg Val Asp His Leu Leu Ile Arg Val
                125                     130                     135
Met Leu Tyr Gln Ile Ser Glu Glu Val Ser Arg Ser Glu Leu Arg
                140                     145                     150
Ser Phe Lys Phe Leu Leu Gln Glu Glu Ile Ser Lys Cys Lys Leu
                155                     160                     165
Asp Asp Asp Met Asn Leu Leu Asp Ile Phe Ile Glu Met Glu Lys
                170                     175                     180
Arg Val Ile Leu Gly Glu Gly Lys Leu Asp Ile Leu Lys Arg Val
                185                     190                     195
Cys Ala Gln Ile Asn Lys Ser Leu Leu Lys Ile Ile Asn Asp Tyr
                200                     205                     210
Glu Glu Phe Ser Lys Gly Glu Glu Leu Cys Gly Val Met Thr Ile
                215                     220                     225
Ser Asp Ser Pro Arg Glu Gln Asp Ser Glu Ser Gln Thr Leu Asp
                230                     235                     240
Lys Val Tyr Gln Met Lys Ser Lys Pro Arg Gly Tyr Cys Leu Ile
                245                     250                     255
Ile Asn Asn His Asn Phe Ala Lys Ala Arg Glu Lys Val Pro Lys
                260                     265                     270
Leu His Ser Ile Arg Asp Arg Asn Gly Thr His Leu Asp Ala Gly
                275                     280                     285
Thr Val Glu Pro Lys Arg Glu Lys
                290

<210> 55
<211> 511
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510233CD1

<400> 55
Met Asp Phe Ser Arg Asn Leu Tyr Asp Ile Gly Glu Gln Leu Asp
  1               5                      10                      15
Ser Glu Asp Leu Ala Ser Leu Lys Phe Leu Ser Leu Asp Tyr Ile
                 20                      25                      30
Pro Gln Arg Lys Gln Glu Pro Ile Lys Asp Ala Leu Met Leu Phe
                 35                      40                      45
Gln Arg Leu Gln Glu Lys Arg Met Leu Glu Glu Ser Asn Leu Ser
                 50                      55                      60
Phe Leu Lys Glu Leu Leu Phe Arg Ile Asn Arg Leu Asp Leu Leu
                 65                      70                      75
Ile Thr Tyr Leu Asn Thr Arg Lys Glu Glu Met Glu Arg Glu Leu
                 80                      85                      90
Gln Thr Pro Gly Arg Ala Gln Ile Ser Ala Tyr Arg Phe His Phe
                 95                     100                     105
Cys Arg Met Ser Trp Ala Glu Ala Asn Ser Gln Cys Gln Thr Gln
                110                     115                     120
Ser Val Pro Phe Trp Arg Arg Val Asp His Leu Leu Ile Arg Val
                125                     130                     135
Met Leu Tyr Gln Ile Ser Glu Glu Val Ser Arg Ser Glu Leu Arg
                140                     145                     150
Ser Phe Lys Phe Leu Leu Gln Glu Glu Ile Ser Lys Cys Lys Leu
                155                     160                     165
Asp Asp Asp Met Asn Leu Leu Asp Ile Phe Ile Glu Met Glu Lys
                170                     175                     180
Arg Val Ile Leu Gly Glu Gly Lys Leu Asp Ile Leu Lys Arg Val
                185                     190                     195
Cys Ala Gln Ile Asn Lys Ser Leu Leu Lys Ile Ile Asn Asp Tyr
```

```
                200                 205                 210
Glu Glu Phe Ser Lys Glu Arg Ser Ser Ser Leu Glu Gly Ser Pro
                215                 220                 225
Asp Glu Phe Ser Asn Gly Glu Glu Leu Cys Gly Val Met Thr Ile
                230                 235                 240
Ser Asp Ser Pro Arg Glu Gln Asp Ser Glu Ser Gln Thr Leu Asp
                245                 250                 255
Lys Val Tyr Gln Met Lys Ser Lys Pro Arg Gly Tyr Cys Leu Ile
                260                 265                 270
Ile Asn Asn His Asn Phe Ala Lys Ala Arg Glu Lys Val Pro Lys
                275                 280                 285
Leu His Ser Ile Arg Asp Arg Asn Gly Thr His Leu Asp Ala Gly
                290                 295                 300
Ala Leu Thr Thr Thr Phe Glu Glu Leu His Phe Glu Ile Lys Pro
                305                 310                 315
His Asp Asp Cys Thr Val Glu Gln Ile Tyr Glu Ile Leu Lys Ile
                320                 325                 330
Tyr Gln Leu Met Asp His Ser Asn Met Asp Cys Phe Ile Cys Cys
                335                 340                 345
Ile Leu Ser His Gly Asp Lys Gly Ile Ile Tyr Gly Thr Asp Gly
                350                 355                 360
Gln Glu Ala Pro Ile Tyr Glu Leu Thr Ser Gln Phe Thr Gly Leu
                365                 370                 375
Lys Cys Pro Ser Leu Ala Gly Lys Pro Lys Val Phe Phe Ile Gln
                380                 385                 390
Ala Cys Gln Gly Asp Asn Tyr Gln Lys Gly Ile Pro Val Glu Thr
                395                 400                 405
Asp Ser Glu Glu Gln Pro Tyr Leu Glu Met Asp Leu Ser Ser Pro
                410                 415                 420
Gln Thr Arg Tyr Ile Pro Asp Glu Ala Asp Phe Leu Leu Gly Met
                425                 430                 435
Ala Thr Val Asn Asn Cys Val Ser Tyr Arg Asn Pro Ala Glu Gly
                440                 445                 450
Thr Trp Tyr Ile Gln Ser Leu Cys Gln Ser Leu Arg Glu Arg Cys
                455                 460                 465
Pro Arg Gly Asp Asp Ile Leu Thr Ile Leu Thr Glu Val Asn Tyr
                470                 475                 480
Glu Val Ser Asn Lys Asp Asp Lys Lys Asn Met Gly Lys Gln Met
                485                 490                 495
Pro Gln Pro Thr Phe Thr Leu Arg Lys Lys Leu Val Phe Pro Ser
                500                 505                 510
Asp


<210> 56
<211> 429
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510304CD1

<400> 56
Met Ala Gln Lys Glu Gly Gly Arg Thr Val Pro Cys Cys Ser Arg
 1               5                  10                  15
Pro Lys Val Ala Ala Leu Thr Ala Gly Thr Leu Leu Leu Leu Thr
                20                  25                  30
Ala Ile Gly Ala Ala Ser Trp Ala Ile Gly Glu Ser Gly Ser Cys
                35                  40                  45
Ala Ser Asp Leu Pro Trp Pro Leu Gln Pro Pro Ala Cys Leu Pro
                50                  55                  60
Ser Pro Leu Pro Ser Leu Gln Trp Leu Phe Ser Ser Gly Val Thr
```

```
                65                  70                  75
Arg Ser Arg Cys Thr Gln Cys Lys Thr Trp Arg Leu Leu Cys Ser
                80                  85                  90
Ser Arg Ser Asn Ala Arg Val Ala Gly Leu Ser Cys Glu Glu Met
                95                  100                 105
Gly Phe Leu Arg Ala Leu Thr His Ser Glu Leu Asp Val Arg Thr
                110                 115                 120
Ala Gly Ala Asn Gly Thr Ser Gly Phe Phe Cys Val Asp Glu Gly
                125                 130                 135
Arg Leu Pro His Thr Gln Arg Leu Leu Glu Val Ile Ser Val Cys
                140                 145                 150
Asp Cys Pro Arg Gly Arg Phe Leu Ala Ala Ile Cys Gln Asp Cys
                155                 160                 165
Gly Arg Arg Lys Leu Pro Val Asp Arg Ile Val Gly Gly Arg Asp
                170                 175                 180
Thr Ser Leu Gly Arg Trp Pro Trp Gln Val Ser Leu Arg Tyr Asp
                185                 190                 195
Gly Ala His Leu Cys Gly Gly Ser Leu Leu Ser Gly Asp Trp Val
                200                 205                 210
Leu Thr Ala Ala His Cys Phe Pro Glu Arg Asn Arg Val Leu Ser
                215                 220                 225
Arg Trp Arg Val Phe Ala Gly Ala Val Ala Gln Ala Ser Pro His
                230                 235                 240
Gly Leu Gln Leu Gly Val Gln Ala Val Val Tyr His Gly Gly Tyr
                245                 250                 255
Leu Pro Phe Arg Asp Pro Asn Ser Glu Glu Asn Ser Asn Asp Ile
                260                 265                 270
Ala Leu Val His Leu Ser Ser Pro Leu Pro Leu Thr Glu Tyr Ile
                275                 280                 285
Gln Pro Val Cys Leu Pro Ala Ala Gly Gln Ala Leu Val Asp Gly
                290                 295                 300
Lys Ile Cys Thr Val Thr Gly Trp Gly Asn Thr Gln Tyr Tyr Gly
                305                 310                 315
Gln Gln Ala Gly Val Leu Gln Glu Ala Arg Val Pro Ile Ile Ser
                320                 325                 330
Asn Asp Val Cys Asn Gly Ala Asp Phe Tyr Gly Asn Gln Ile Lys
                335                 340                 345
Pro Lys Met Phe Cys Ala Gly Tyr Pro Glu Gly Gly Ile Asp Ala
                350                 355                 360
Cys Gln Gly Asp Ser Gly Gly Pro Phe Val Cys Glu Asp Ser Ile
                365                 370                 375
Ser Arg Thr Pro Arg Trp Arg Leu Cys Gly Ile Val Ser Trp Gly
                380                 385                 390
Thr Gly Cys Ala Leu Ala Gln Lys Pro Gly Val Tyr Thr Lys Val
                395                 400                 405
Ser Asp Phe Arg Glu Trp Ile Phe Gln Ala Ile Lys Thr His Ser
                410                 415                 420
Glu Ala Ser Gly Met Val Thr Gln Leu
                425
```

<210> 57
<211> 412
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510461CD1

<400> 57

```
Met Ala Asn Val Gly Leu Gln Phe Gln Ala Ser Ala Gly Asp Ser
  1             5                   10                  15
Asp Pro Gln Ser Arg Pro Leu Leu Leu Leu Gly Gln Leu His His
```

20 25 30
Leu His Arg Val Pro Trp Ser His Val Arg Gly Lys Leu Gln Pro
35 40 45
Arg Val Thr Glu Glu Leu Trp Gln Ala Ala Leu Ser Thr Leu Asn
50 55 60
Pro Asn Pro Thr Asp Ser Cys Pro Leu Tyr Leu Asn Tyr Ala Thr
65 70 75
Val Ala Ala Leu Pro Cys Arg Val Ser Arg His Asn Ser Pro Ser
80 85 90
Ala Ala His Phe Ile Thr Arg Leu Val Arg Thr Cys Leu Pro Pro
95 100 105
Gly Ala His Arg Cys Ile Val Met Val Cys Glu Gln Pro Glu Val
110 115 120
Phe Ala Ser Ala Cys Ala Leu Ala Arg Ala Phe Pro Leu Phe Thr
125 130 135
His Arg Ser Gly Ala Ser Arg Arg Leu Glu Lys Lys Thr Val Thr
140 145 150
Val Glu Phe Phe Leu Val Gly Gln Asp Asn Gly Pro Val Glu Val
155 160 165
Ser Thr Leu Gln Cys Leu Ala Asn Ala Thr Asp Gly Val Arg Leu
170 175 180
Ala Ala Arg Ile Val Asp Thr Pro Cys Asn Glu Met Asn Thr Asp
185 190 195
Thr Phe Leu Glu Glu Ile Asn Lys Val Gly Lys Glu Leu Gly Ile
200 205 210
Ile Pro Thr Ile Ile Arg Asp Glu Glu Leu Lys Thr Arg Gly Phe
215 220 225
Gly Gly Ile Tyr Gly Val Gly Lys Ala Ala Leu His Pro Pro Ala
230 235 240
Leu Ala Val Leu Ser His Thr Pro Asp Gly Ala Thr Gln Thr Ile
245 250 255
Ala Trp Val Gly Lys Gly Ile Val Tyr Asp Thr Gly Gly Leu Ser
260 265 270
Ile Lys Gly Lys Thr Thr Met Pro Gly Met Lys Arg Asp Cys Gly
275 280 285
Gly Ala Ala Ala Val Leu Gly Ala Phe Arg Ala Ala Ile Lys Gln
290 295 300
Gly Phe Lys Asp Asn Leu His Ala Val Phe Cys Leu Ala Glu Asn
305 310 315
Ser Val Gly Pro Asn Ala Thr Arg Pro Asp Asp Ile His Leu Leu
320 325 330
Tyr Ser Gly Lys Thr Val Glu Ile Asn Asn Thr Asp Ala Glu Gly
335 340 345
Arg Leu Val Leu Ala Asp Gly Val Ser Tyr Ala Cys Lys Asp Leu
350 355 360
Gly Ala Asp Ile Ile Leu Asp Met Ala Thr Leu Thr Gly Ala Gln
365 370 375
Val Ser Ala Pro Trp Ile His Pro Leu Ala Val Val Pro Gly Asn
380 385 390
Pro Thr Pro Leu Leu Thr Ser Arg Trp Gly Arg Gln Arg Val Gly
395 400 405
Ala Ala Val Arg Lys Phe Ser
410

<210> 58
<211> 123
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510392CD1

<400> 58
Met Ala Val Val Pro Leu Leu Leu Leu Gly Gly Leu Trp Ser Ala
1 5 10 15
Val Gly Ala Ser Ser Leu Gly Val Val Thr Cys Gly Ser Val Val
20 25 30
Lys Leu Leu Asn Thr Arg His Asn Val Arg Leu His Ser His Asp
35 40 45
Val Arg Tyr Gly Ser Gly Ser Gly Gln Gln Ser Val Thr Gly Val
50 55 60
Thr Ser Val Asp Asp Ser Asn Ser Tyr Trp Arg Ile Arg Gly Lys
65 70 75
Ser Ala Thr Val Cys Glu Arg Gly Thr Pro Ile Lys Cys Gly Gln
80 85 90
Pro Ile Arg Leu Thr His Val Asn Thr Gly Arg Asn Leu His Ser
95 100 105
His His Phe Thr Ser Pro Leu Ser Gly Asn Gln Leu Leu Cys Lys
110 115 120
Val Ile Leu

<210> 59
<211> 1389
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7313196CB1

<400> 59
ggacggcagg accaagaagc agctcgccat ggtggaagga accatgaact ggcatctccc 60
cctcttcctc ttggcctctg tgacgctgcc ttccatctgc tcccacttca atcctctgtc 120
tctcgaggaa ctaggctcca acacggggat ccaggttttc aatcagattg tgaagtcgag 180
gcctcatgac aacatcgtga tctctcccca tgggattgcg tcggtcctgg ggatgcttca 240
gctgggggcg gacggcagga ccaagaagca gctcgccatg gtgatgagat acggcgtaaa 300
tgatatgatt gacaatctgc tgtccccaga tcttattgat ggtgtgctca ccagactggt 360
cctcgtcaac gcagtgtatt tcaagggtct gtggaaatca cggttccaac ccgagaacac 420
aaagaaacgc actttcgtgg cagccgacgg gaaatcctat caagtgccaa tgctggccca 480
gctctccgtg ttccggtgtg ggtcgacaag tgcccccaat gatttatggt acaacttcat 540
tgaactgccc taccacgggg aaagcatcag catgctgatt gcactgccga ctgagagctc 600
cactccgctg tctgccatca tcccacacat cagcaccaag accatagaca gctggatgag 660
catcatggtg cccaagaggg tgcaggtgat cctgcccaag ttcacagctg tagcacaaac 720
agatttgaag gagccgctga aagttcttgg cattactgac atgtttgatt catcaaaggc 780
aaattttgca aaaataacaa ggtcagaaaa cctccatgtt tctcatatct tgcaaaaagc 840
aaaaattgaa gtcagtgaag atggtaccaa agcttcagca gcaacaactg caattctcat 900
tgcaagatca tcgcctcccct ggtttatagt agacagacct tttctgtttt tcatccgaca 960
taatcctaca ggtgctgtgt tattcatggg gcagataaac aaaccctgaa gagtatacaa 1020
gagaaaccat gcaaagcaac gactactttg ctacgaagaa agactccttt cctgcatctc 1080
gcttcttttt caaaactagt tcttaggaac agactcgatg caagtgtttc tgttctggga 1140
ggtattggag ggaaaaaaca agcaggatgg ctggaacact gtactgagga atgaatagaa 1200
aggcttccag atgtctaaaa gattctttaa actactgaac tgttacctag ggtaacaacc 1260
cctgtgagta tttgctgttt gtcgggtcag gaattttgtt ttgtttgtct atatgtgcgg 1320
ctttacagag aatttatcag tgtacagacc aatatgttta tgtagcttta cttttatgaa 1380
aaaaattat 1389

<210> 60
<211> 4131
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6465289CB1

<400> 60
ggcggcggcg gcggagccct gggtcggtgt ctgcgcgctg gtgtctgagg cccaggctga 60
ggcctccgct attgctggag cgcaggcggc ggagaggatg actgccgctg ccattctctc 120
ttgagctagc gagccgccgc caccctccac cctcccccgg cagggcggag aggagcggcc 180
ggagtcagcg atggtgcccg gcgaggagaa ccaactggtc ccgaaagaga tagaaaatgc 240
tgctgaagaa cctagagtct tatgtattat acaagatact actaattcaa agacagtgaa 300
tgaacggatc actttaaatt taccagcatc tactccagtc agaaagctct ttgaagatgt 360
ggccaacaaa gtaggctaca taaatggaac ctttgacttg gtgtggggaa atggaatcaa 420
tactgctgat atggcaccac tggatcatac cagtgacaag tcacttctcg acgctaattt 480
tgagccagga aagaagaact ttctgcattt gacagataaa gatggtgaac aacctcaaat 540
actgctggag gattccagtg ctggggaaga cagtgttcat gacaggttta taggtccgct 600
tccaagagaa ggttctgtgg gttctaccag tgattatgtc agccaaagct actcctactc 660
atctattttg aataaatcag aaactggata tgtgggacta gtaaaccaag caatgacttg 720
ctatttgaat agccttttgc aaacactttt tatgactcct gaatttagga atgcattata 780
taagtgggaa tttgaagaat ctgaagaaga tccagtgaca agtattccat accaacttca 840
aaggcttttt gttttgttac aaaccagcaa aaagagagca attgaaacca cagatgttac 900
aaggagcttt ggatgggata gtagtgaggc ttggcagcag catgatgtac aagaactatg 960
cagagtcatg tttgatgctt tggaacagaa atggaagcaa acagaacagg ctgatcttat 1020
aaatgagcta tatcaaggca agctgaagga ctacgtgaga tgtctggaat gtggttatga 1080
gggctggcga atcgacacat atcttgatat tccattggtc atccgacctt atgggtccag 1140
ccaagcattt gctagtgtgg aagaagcatt gcatgcattt attcagccag agattctgga 1200
tggcccaaat cagtattttt gtgaacgttg taagaagaag tgtgatgcac ggaagggcct 1260
tcggtttttg cattttcctt atctgctgac cttacagctg aaaagattcg attttgatta 1320
tacaaccatg cataggatta aactgaatga tcgaatgaca tttcccgagg aactagatat 1380
gagtactttt attgatgttg aagatgagaa atctcctcag actgaaagtt gcactgacag 1440
tggagcagaa aatgaaggta gttgtcacag tgatcagatg agcaacgatt tctccaatga 1500
tgatggtgtt gatgaaggaa tctgtcttga aaccaatagt ggaactgaaa agatctcaaa 1560
atctggactt gaaaagaatt ccttgatcta tgaacttttc tctgttatgg ttcattctgg 1620
gagcgctgct ggtggtcatt attatgcatg tataaagtca ttcagtgatg agcagtggta 1680
cagcttcaat gatcaacatg tcagcaggat aacacaagag gacattaaga aaacacatgg 1740
tggatcttca ggaagcagag gatattattc tagtgctttc gcaagttcca caaatgcata 1800
tatgctgatc tatagactga aggatccagc cagaaatgca aaatttctag aagtggatga 1860
atacccagaa catattaaaa acttggtgca gaaagagaga gagttggaag aacaagaaaa 1920
gagacaacga gaaattgagc gcaatacatg caagataaaa ttattctgtt tgcatcctac 1980
aaaacaagta atgatggaaa ataaattgga ggttcataag gataagacat taaaggaagc 2040
agtagaaatg gcttataaga tgatggattt agaagaggta atacccctgg attgctgtcg 2100
ccttgttaaa tatgatgagt ttcatgatta tctagaacgg tcatatgaag gagaagaaga 2160
tacaccaatg gggcttctac taggtggcgt caagtcaaca tatatgtttg atctgctgtt 2220
ggagacgaga aagcctgatc aggttttcca atcttataaa cctggagaag tgatggtgaa 2280
agttcatgtt gttgatctaa aggcagaatc tgtagctgct cctataactg ttcgtgctta 2340
cttaaatcag acagttacag aattcaaaca actgatttca aaggccatcc atttacctgc 2400
tgaaacaatg agaatagtgc tggaacgctg ctacaatgat ttgcgtcttc tcagtgtctc 2460
cagtaaaacc ctgaaagctg aaggattttt tagaagtaac aaggtgtttg ttgaaagctc 2520
cgagactttg gattaccaga tggcctttgc agactctcat ttatggaaac tcctggatcg 2580
gcatgcaaat acaatcagat tatttgtttt gctacctgaa caatccccag tatcttattc 2640
caaaaggaca gcataccaga aagctggagg cgattctggt aatgtggatg atgactgtga 2700
aagagtcaaa ggacctgtag gaagcctaaa gtctgtggaa gctattctag aagaaagcac 2760
tgaaaaactc aaaagcttgt cactgcagca acagcaggat ggagataatg gggacagcag 2820
caaaagtact gagacaagtg actttgaaaa catcgaatca cctctcaatg agagggactc 2880
ttcagcatca gtggataata gagaacttga acagcatatt cagacttctg atccagaaaa 2940
ttttcagtct gaagaacgat cagactcaga tgtgaataat gacaggagta caagttcagt 3000
ggacagtgat attcttagct ccagtcatag cagtgatact ttgtgcaatg cagacaatgc 3060
tcagatccct ttggctaatg gacttgactc tcacagtatc acaagtagta gaagaacgaa 3120
agcaaatgaa gggaaaaaag aaacatggga tacagcagaa gaagactctg gaactgatag 3180
tgaatatgat gagagtggca agagtagggg agaaatgcag tacatgtatt tcaaagctga 3240
accttatgct gcagatgaag gttctgggga aggacataaa tggttgatgg tgcatgttga 3300
taaaagaatt actctggcag ctttcaaaca acatttagag ccctttgttg gagttttgtc 3360
ctctcacttc aaggtctttc gagtgtatgc cagcaatcaa gagtttgaga gcgtccggct 3420
gaatgagaca ctttcatcat ttctgatga caataagatt acaattagac tggggagagc 3480
acttaaaaaa ggagaataca gagttaaagt ataccagctt ttggtcaatg aacaagagcc 3540
atgcaagttt ctgctagatg ctgtgtttgc taaaggaatg actgtacggc aatcaaaaga 3600
ggaattaatt cctcagctca gggagcaatg tggtttagag ctcagtattg acaggtttcg 3660
tctaaggaaa aaaacatgga agaatcctgg cactgtcttt ttggattatc atatttatga 3720

agaagatatt aatatttcca gcaactggga ggttttcctt gaagttcttg atggggtaga 3780
gaagatgaag tccatgtcac agcttgcagt tttgtcaaga cggtggaagc cttcagagat 3840
gaagttggat cccttccagg aggttgtatt ggaaagcagt agtgtggacg aattgcgaga 3900
gaagcttagt gaaatcagtg ggattccttt ggatgatatt gaatttgcta agggtagagg 3960
aacatttccc tgtgatattt ctgtccttga tattcatcaa gatttagact ggaatcctaa 4020
agtttctacc ctgaatgtct ggcctcttta tatctgtgat gatggtgcgg tcatatttta 4080
tagggataaa acagaagaat taatggagat tgacagatga ctaaagctat g 4131

<210> 61
<211> 3249
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506357CB1

<400> 61
aggcgcgcca gaaagaaggg agaccagcag cgtcccttat tcgcttggcc ttggttcctg 60
tttgcactgg ctacagcagg gcactggccc ctactgtcac cgccacctac acaaagaccc 120
tatctctgag cgctgcagcc tactgttcag ccccaggttt gaggatggat gccctggacg 180
cttcgaagct actggatgag gagctgtatt caagacagct gtatgtgctg ggctcacctg 240
ccatgcagag gattcaggga gccagggtcc tggtgtcagg cctgcagggc ctgggggccg 300
aggtggccaa gaacttggtt ctgatgggtg tgggcagcct cactctgcat gatccccacc 360
ccacctgctg gtccgacctg gctgcccagt ttctcctctc agagcaggac ttggaaagga 420
gcagagccga ggcctctcaa gagctcttgg ctcagctcaa cagagctgtc caggtcgtcg 480
tgcacacggg tgacatcact gaggacctgc tgttggactt ccaggtggtg gtgctgactg 540
ctgcaaagct ggaggagcag ctgaaggtgg gcaccttgtg tcataagcat ggagtttgct 600
ttctggcggc tgacacccgg ggcctcgtgg ggcagttgtt ctgtgacttt ggtgaggact 660
tcactgtgca ggaccccaca gaggcagaac ccctgacagc tgccatccag cacatctccc 720
agggctcccc tggcattctc actctgagga aaggggccaa tacccactac ttccgtgatg 780
gagacttggt gactttctcg ggaattgagg gaatggttga gctcaacgac tgtgatcccc 840
ggtctatcca cgtgcgggag gatgggtccc tggagattgg agacacaaca actttctctc 900
ggtacttgcg tggtggggct atcactgaag tcaagagacc caagactgtg agacataagt 960
ccctggacac agccctgctc cagccccatg tggtggccca gagctcccag gaagttcacc 1020
atgcccactg cctgcatcag gccttctgtg cactgcacaa gttccagcac ctccatggcc 1080
ggccacccca gccctgggat cctgttgatg cagagactgt ggtgggcctg gcccgggacc 1140
tggaaccact gaagcggaca gaggaagagc cactggaaga gccactggat gaggccctag 1200
tgcggacagt cgccctaagc agtgcaggtg tcttgagccc tatggtggcc atgctgggtg 1260
cagtagctgc ccaggaagtg ctgaaggcaa tctccaggaa gttcatgcct ctggaccagt 1320
ggctttactt tgatgccctc gattgtcttc cggaagatgg ggagctcctt cccagtcctg 1380
aggactgtgc cctgagaggc agccgctatg atgggcaaat tgcagtgttt ggggctggtt 1440
ttcaggagaa actgagacgc cagcactacc tcctggtggg cgctggtgcc attggttgtg 1500
agctgctcaa agtctttgcc ctagtgggac tgggggccgg gaacagcggg ggcttgactg 1560
ttgttgacat ggaccacata gagcgctcca atctcagccg tcagttcctc ttcaggtccc 1620
aggacgttgg tagacccaag gcagaggtgg ctgcagcagc tgcccggggc ctgaacccag 1680
acttacaggt gatcccgctc acctacccac tggatcccac cacagagcac atctatgggg 1740
ataacttttt ctcccgtgtg gatggtgtgg ctgctgccct ggacagtttc caggcccggc 1800
gctatgtggc tgctcgttgc acccactatc tgaagccact gctggaggca ggcacatcgg 1860
gcacctgggg cagtgctaca gtattcatgc cacatgtgac tgaggcctac agagcccctg 1920
cctcagctgc agcttctgag gatgcccccct accctgtctg taccgtgcgg tacttcccta 1980
gcacagccga gcacaccctg cagtgggccc ggcatgagtt tgaagaactc ttccgactgt 2040
ctgcagagac catcaaccac caccaacagg cacacacttc cctggcagac atggatgagc 2100
cacagacact caccttactg aagccagtgc ttggggtcct gagagtgcgt ccacagaact 2160
ggcaagactg tgtggcgtgg gctcttggcc actggaaact ctgctttcat tatggcatca 2220
aacagctgct gaggcacttc ccacctaata aagacacaca cctcctctac gtactggcag 2280
ctgccaacct gtatgcccag atgcatgggc tgcctggctc acaggactgg actgcactca 2340
gggagctgct gaagctgctg ccacagcctg acccccaaca gatggccccc atctttgcta 2400
gtaatctaga gctggcttcg gcttctgctg agtttggccc tgagcagcag aaggaactga 2460
acaaagccct ggaagtctgg agtgtgggcc ctcccctgaa gcctctgatg tttgagaagg 2520
atgatgacag caacttccat gtggactttg tggtagcggc agctagcctg agatgtcaga 2580
actacgggat tccaccggtc aaccgtgccc agagcaagcg aattgtgggc cagattatcc 2640
cagccattgc caccactaca gcagctgtgg caggcctgtt gggcctggag ctgtataagg 2700

tggtgagtgg gccacggcct cgtagtgcct ttcgccacag ctacctacat ctggctgaaa 2760
actacctcat ccgctatatg ccttttgccc cagccatcca gacgttccat cacctgaagt 2820
ggacctcttg ggaccgtctg aaggtaccag ctgggcagcc tgagaggacc ctggagtcgc 2880
tgctggctca tcttcaggag cagcacgggt tgagggtgag gatcctgctg cacggctcag 2940
ccctgctcta tgcggccgga tggtcacctg aaaagcaggc ccagcacctg cccctcaggg 3000
tgacagaact ggttcagcag ctgacaggcc aggcacctgc tcctgggcag cgggtgttgg 3060
tgctagagct gagctgtgag ggtgacgacg aggacactgc cttcccacct ctgcactatg 3120
agctgtgaca aggcagccac cctgtcacct agctcaatgg agccccggat cccaagccct 3180
gcattgtaag cccacagtag gcactcaata aatgcttgtt aaaggaaaaa aaaaaaaaaa 3240
aaaaaaaaa 3249

<210> 62
<211> 986
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6878857CB1

<400> 62
ggagatagat gagcgtacgt aagctcggaa tcggtcgagc atgaatctac ttctgatcct 60
tacctttgtt gcagctgctg ttgctgcccc ctttgatgat gatgacaaga tcgttggggg 120
ctacatctgt gaggagaatt ctgtccccta ccaggtgtcc ttgaattctg gctacaagtc 180
ccgcatccag gtgagactgg gagagcacaa catcgaagtc ctggagggga atgaacagtt 240
catcaatgca gccaagatca tccgccaccc caaatacaac agccggactc tggacaatga 300
catcctgctg atcaagctct cctcacctgc cgtcatcaat tcccgcgtgt ccgccatctc 360
tctgcccact gcccctccag ctgctggcac cgagtccctc atctccggct ggggcaacac 420
tctgagttct ggtgccgact acccagacga gctgcagtgc ctggatgctc ctgtgctgag 480
ccaggctgag tgtgaagcct cctaccctgg aaagattacc aacaacatgt tctgtgtggg 540
cttcctcgag ggaggcaagg attcctgcca gggtgattct ggtggccctg tggtctccaa 600
tggagagctc caaggaattg tctcctgggg ctatggctgt gcccagaaga acaggcctgg 660
agtctacacc aaggtctaca actatgtgga ctggattaag gacaccatag ctgccaacag 720
ctaaagcccc tggtccctct gcagtctcta taccaataaa gtgaccctgc tctcacaaaa 780
aaaaaaaaaa ggcgccgcca ataattaatc tcttcaaccc gaaataattc ccgaaccgtt 840
accttcgcct tgcttggaat ttcattttag ttatcgttac cgtccactcg agggggcccc 900
gtacccatcg cctatattat gtattaccgt aattggcgtg tttaacttct gttgaactgg 960
ttacattatc ctgtatcttt caaaaa 986

<210> 63
<211> 3665
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506021CB1

<400> 63
tactgccttc aagagatttt tctctctgca gtattcatgg tcagaattct ttgctgatga 60
caacagcttc tctgcttgga gtctcaggag agtctgtgag gaggaaatga agtgtttagc 120
gtccgttctg gtcactttag aggacaaagt actaacattt tgagtagcca cattgaactt 180
ggtcttagca cataaggggt tcatccttga tccacgtttt ctggatttga agagctttct 240
ttctaccctg gaaatgtctc aactgtctat tccaagtaac tggccttccc agcttagctt 300
tcagaatgca caagtgttgg tgtaaatata acttctttaa gcctccaaac ccagtattcc 360
atttctcaga aaaggctaag tggattcctt tcctccatag aattttcctc tgttttttca 420
tttttctgat gcataatgag atctggatgc atgctgctac cctcgttcct ggcagctgcc 480
ccgccgctgg ccggcgtcgc ggcagccgct tcagaggaag ctcggcggct gtggccgcgg 540
cgaacaggcc ggacgcctcg tcgctggcgg ggcttccctt ggagctgacg aaggcggggt 600
ccgcggtcca ggctgctgcc gcgacggggc cggcggcggg gcagctgcca ggaacgaggg 660
tagcagctgc atccagatct cattatgcat cagaaaaatg aaaaaacaga ggaaaattct 720
atggaggaaa ggaatccact tagccttttc tgagaaatgg aatactgggt ttggaggctt 780
taagaagttt tattttcacc aacacttgtg cattctgaaa gctaagctgg gaaggccagt 840

tacttggaat agacagttga gacatttcca gggtagaaag aaagctcttc aaatccagaa 900
aacgtggatc aaggatgaac ccctttgtgc taagaccaag ttcaatgtgg ctactcaaaa 960
tgttagtact ttgtcctcta aagtgaaaag aaaggacgct aaacacttca tttcctcctc 1020
aaagactctc ctgagactcc aagcagagaa gctgttgtca tcagcaaaga attctgacca 1080
tgaatactgc agagagaaaa atctcttgaa ggcagttact gactttccat caaatagtgc 1140
tttaggtcag gccaatggtc acagacctag gacagaccca caaccttctg actttcccat 1200
gaagttcaat ggggagagcc aaagtccagt tgagagtggc acgattgtgg tcaccttgaa 1260
caaccataag agaaagggct tttgttaccg ctgctgccaa gggccggagc accacaggaa 1320
tgggggaccc ttgattccaa aaaagttcca acttaaccaa catagaagga taaaattatc 1380
tcctcttatg atgtatgaga aattatccat gattagattt cggtacagga ttctcagatc 1440
ccagcacttc agaaccaaaa gcaaggtttg caagctaaga aaagcccagc gaagctgggt 1500
acagaaagtc actggggacc atcaagagac ccgtagggag aacggtgagg gtggcagttg 1560
cagcccattt ccttccccag aacctaaaga cccttcttgt cggcatcagc cgtactttcc 1620
agatatggac agcagtgctg tggtgaaggg gacgaactct catgtgcctg attgccacac 1680
taaaggaagc tctttcttgg gcaaggagct tagtttagat gaagcattcc ctgaccaaca 1740
gaatggcagt gccacaaacg cctgggacca gtcatcctgt tcttctccta agtgggagtg 1800
tacagagctg attcatgaca tccccttacc agaacatcgt tctaatacca tgttcatttc 1860
agaaactgaa agagaaatta tgactctggg tcaggaaaat cagacaagtt ctgtcagtga 1920
tgacagagta aaactgtcag tgtctggagc agatacatct gtgagtagcg tagatgggcc 1980
tgtgtcccaa aaggctgttc aaaatgagaa ctcataccag atggaggagg atggatctct 2040
caagcagagc attcttagtt ctgagttgct ggaccacccct tactgtaaaa gtccactgga 2100
ggctcccttg gtgtgcagtg gactcaaact agaaaatcaa gtaggaggtg gaaagaacag 2160
tcagaaagcc tctccagtgg atgatgaaca gctgtcagtc tgtctttctg gattcctaga 2220
tgaggttatg aagaagtatg gcagtttggt tccactcagt gaaaaagaag tccttggaag 2280
attaaaagat gtctttaatg aagacttttc taatagaaaa ccatttatca atagggaaat 2340
aacaaactat cgggccagac atcaaaaatg taacttccgt atcttctata ataaacacat 2400
gctggatatg gacgacctgg cgactctgga tggtcagaac tggctgaatg accaggtcat 2460
taatatgtat ggtgagctga taatggatgc agtcccagac aaagttcact tcttcaacag 2520
ctttttttcat agacagctgg taaccaaagg atataatgga gtaaaaagat ggactaaaaa 2580
ggtggatttg tttaaaaaga gtcttctgtt gattcctatt cacctggaag tccactggtc 2640
tctcattact gtgacactct ctaatcgaat tatttcattt tatgattccc aaggcattca 2700
ttttaagttt tgtgtagagt gtattccaca acagaaaaac gacagtgact gtggagtctt 2760
tgtgctccag tactgcaagt gcctcgcctt agagcagcct ttccagtttt cacaagaaga 2820
catgccccga gtgcggaaga ggatttacaa ggagctatgt gagtgccggc tcatggactg 2880
aaactcagca gggactctgg ggaagtctgac caagttggag cagatggttt gttacttgaa 2940
tctccaaaca cttagttgaa tttttacaga tatttcagat cagtggtgtt gggccactat 3000
tgttacctca aatttatttt ttgcccttat tcatttctcc agctaccatg tactattgtt 3060
taatgttcag tttggtttca tttttaattt tatggttctg tgcgtccccc atatttaata 3120
tttattattc aaacgcatgc atatagacag agcatgcagt gaagagtatt aaaaaaaaaa 3180
gcttagtaga tttggtgcag cttttgaaac ttagttagac gtgaactgaa tacaggtttc 3240
aaatttactc ccagaaccta aaaatgcaag atgtttttga tacaacataa ctctgagaat 3300
agtaagtgtt ccctggggca ttaagggtag ctgggggtgg ttttgacaaa tccagtcctg 3360
ttttacttta ccagcggcaa ctttcaccaa cttcccctct ccaagtgagt cttagagagt 3420
gcagtccatt ccttttgaag ggtgagatgg aagtggtcgt aaactgactg gtgtcttctg 3480
tttctggagg cacacttgta agcacagtgg ctgctttggg aggagtaagg tgtgagaaaa 3540
agcaaccttg gaggccagta acaatgacag atttcaatcg tggttttagg aattataata 3600
cgtggcatac atctcataaa ggcttttgct gggaaaaaaa aaaaaaaaaa aaaaaaaaga 3660
tcttt 3665

<210> 64
<211> 1440
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503356CB1

<400> 64
cggacgcgtg ggcggacgcg tgggtgccca ggcgcttaaa gaagcaaaat ctcttgtgca 60
ggagcagcag agactcctca ggaagactca ctggactgta cccaccacct gccatgtctc 120
tgtggccacc tttccgatgc agatggaagc tggcgccaag gtactctagg agggcgtctc 180
cacagcaacc ccaacaggac tttgaggccc tgctggcaga gtgcctgagg aatggctgcc 240

tctttgaaga caccagcttc ccggccaccc tgagctccat cggcagtggc tccctgctgc 300
agaagctgcc accccgcctg cagtggaaga ggcccccgga gctgcacagc aatccccagt 360
tttattttgc caaggccaaa aggctggatc tgtgccaggg gatagtagga gactgctggt 420
tcttggctgc tttgcaagct ctggccttgc accaggacat cctgagccgg gttgttcccc 480
tgaatcagag tttcactgag aagtatgctg gcatcttccg gttctgggct ctctggttcc 540
tatgaagact tgcagtcagg acaggtgtct gaagcccttg tagacttcac tggaggggtg 600
acaatgacca tcaacctggc agaagcccat ggcaacctct gggacatcct catcgaagcc 660
acctacaaca gaaccctcat tggctgccag acccactcag gggagaagat tctggagaat 720
gggctggtgg aaggccatgc ctatactctc acaggaatca ggaaggtgac ctgcaaacat 780
agacctgaat atctcgtcaa gctacggaac ccctggggaa aggtggaatg gaaaggagac 840
tggagtgaca gttcaagtaa atgggagctg ctgagcccca aggagaagat tctgcttctg 900
aggaaagaca atgacggaga attctggatg acgctgcagg actttaaaac acatttcgtg 960
ctcctggtta tctgtaaact gaccccaggc ctgttgagcc aggaggcggc ccagaagtgg 1020
acgtacacca tgcgggaggg gagatgggag aagcggagca cagctggtgg ccagaggcag 1080
ttgctgcagg acacattttg gaagaacccg cagttcctgc tgtctgtctg gaggcccgag 1140
gagggcagga gatccctgag gccctgcagc gtgctggtgt ccctgctcca gaagcccagg 1200
cacaggtgcc gcaagcggaa gcctctcctc gccattggct tctacctcta taggatgaac 1260
aagtaccatg atgaccagag gagactgccc cctgagttct tccagagaaa cactcctctg 1320
agccagcctg ataggtttct caaggagaaa gaagtgagtc aggagctgtg tctggaacca 1380
gggacgtacc tcatcgtgcc tgcatattgg aggcccacca gaagtcagag ttcgtcctca 1440

<210> 65
<211> 1064
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509052CB1

<400> 65
gacgccaccg ccgaggaaaa ccgtgtacta ttagccatgg tcaaccccac cgtgttcttc 60
gacattgccg tcgacggcga gcccttgggc cgcgtctcct ttgaggtaag gggcctggat 120
accaagaagt gactgctcat ctaatccata aagctatgtt aacagattgg aggtagtagc 180
attttcatta caagtgacta aaagaacagc tgtttacccc tgatcgtgca gcagtgcttg 240
ctgttcctta gaattttgcc ttctgtttgc agacaaggtc ccaaagacag cagaaaattt 300
tcgtgctctg agcactggag agaaaggatt tggttataag ggttcctgct ttcacagaat 360
tattccaggg tttatgtgtc agggtggtga cttcacacgc cataatggca ctggtggcaa 420
gtccatctat ggggagaaat ttgaagatga gaacttcatc ctaaagcata cgggtcctgg 480
catcttgtcc atggcaaatg ctggacccaa cacaaatggt tcccagtttt tcatctgcac 540
tgccaagact gagtggttgg atggcaagca tgtggtgttt ggcaaagtga aagaaggcat 600
gaatattgtg gaggccatgg agcgctttgg gtccaggaat ggcaagacca gcaagaagat 660
caccattgct gactgtggac aactcgaata agtttgactt gtgttttatc ttaaccacca 720
gatcattcct tctgtagctc aggagagcac ccctccaccc catttgctcg cagtatccta 780
gaatctttgt gctctcgctg cagttccctt tgggttccat gttttccttg ttccctccca 840
tgcctagctg gattgcagag ttaagtttat gattatgaaa taaaaactaa ataacaaaga 900
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 960
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaag 1020
aaaaaaaaaa aaaaaaaggg ggcccccttt agggggcccc cctt 1064

<210> 66
<211> 4497
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503366CB1

<400> 66
cgcggctggg cgcctcggcc atgactgcgg agctgcagca ggacgacgcg gccggcgcgg 60
cagacggcca cggctcgagc tgccaaatgc tgttaaatca actgagagaa atcacaggca 120
ttcaggaccc ttcctttctc catgaagctc tgaaggccag taatggtgac attactcagg 180

cagtcagcct tctcactgat gagagagtta aggagcccag tcaagacact gttgctacag 240
aaccatctga agtagagggg agtgctgcca acaaggaagt attagcaaaa gttatagacc 300
ttactcatga taacaaagat gatcttcagg ctgccattgc tttgagtcta ctggagtctc 360
ccaaaattca agctgatgga agagatctta acaggatgca tgaagcaacc tctgcagaaa 420
ctaaacgctc aaagagaaaa cgctgtgaag tctggggaga aaccccaat cccaatgact 480
ggaggagagt tgatggttgg ccagttgggc tgaaaaatgt tggcaataca tgttggttta 540
gtgctgttat tcagtctctc tttcaattgc ctgaatttcg aagacttgtt ctcagttata 600
gtctgccaca aaatgtactt gaaaattgtc gaagtcatac agaaaagaga aatatcatgt 660
ttatgcaaga gcttcagtat ttgtttgctc taatgatggg atcaaataga aaatttgtag 720
acccgtctgc agccctggat ctattaaagg gagcattccg atcatctgag gaacagcagc 780
aagatgtgag tgaattcaca cacaagctcc tggattggct agaggacgca ttccagctag 840
ctgttaatgt taacagcagt cccaggaaca aatctgaaaa tccaatggtg cagctgttct 900
atggtacttt cctgactgaa ggggttcgtg aaggaaaacc cttttgtaac aatgagacct 960
tcggccagta tcctcttcag gtaaacggtt atcgcaactt agacgagtgt ttggaagggg 1020
ccatggtgga gggtgatgtt gagcttcttc cctccgatca ctcggtgaag tatggacaag 1080
agcgttggtt tacaaagcta cctccagtgt tgacctttga actctcaaga tttgagttta 1140
atcagtccct tgggcagcca gagaaaattc acaataagct ggaatttcct cagattattt 1200
atatggacag gtacatgtac aggagcaagg agcttattcg aaataagaga gagtgtattc 1260
gaaagttgaa ggaggaaata aaaattctgc agcaaaaatt ggaaaggtat gtgaaatatg 1320
gctcaggccc agctcggttc ccgctcccgg acatgctgaa atatgttatt gaatttgcta 1380
gtacaaaacc tgcctcagaa agctgtccac ctgaaagtga cacacatatg acattaccac 1440
tttcttcagt gcactgctcg gtttctgacc agacatccaa ggaaagtaca agtacagaaa 1500
gctcttctca ggatgttgaa agtacctttt cttctcctga agattcttta cccaagtcta 1560
aaccactgac atcttctcgg tcttccatgg aaatgccttc acagccagct ccacgaacag 1620
tcacagatga ggagataaat tttgttaaga cctgtcttca gagatggagg agtgagattg 1680
aacaagatat acaagattta aagacttgta ttgcaagtac tactcagact attgaacaga 1740
tgtactgcga tcctctcctt cgtcagttct tgttcatgaa ggacaagcaa atgctggaca 1800
ctattgggcc tatatctata atcaaccccg acagagctgg ctcaagtaca atgacatctc 1860
tgttactgaa tcttcctggg aagaagttga aagagattcc tatggaggcc tgagaaatgt 1920
tagtgcttac tgtctgatgt acattaatga caaactaccc tacttcaatg cagaggcagc 1980
cccaactgaa tcagatcaaa tgtcagaagt ggaagcccta tctgtggaac tcaagcatta 2040
cattcaggag gataactggc ggtttgagca ggaagtagag gagtgggaag aagagcagtc 2100
ttgcaaaatc cctcaaatgg agtcctccac caactcctca tcacaggact actctacatc 2160
acaagagcct tcagtagcct cttctcatgg ggttcgctgc ttgtcatctg agcatgctgt 2220
gattgtaaag gagcaaactg cccaggctat tgcaaacaca gcccgtgcct atgagaagag 2280
cggtgtagaa gcggcactga gtgaggcatt ccatgaagaa tactccaggc tctatcagct 2340
tgccaaagag acccccacct ctcacagtga tcctcgactt cagcatgtcc ttgtctactt 2400
tttccaaaat gaagcaccca aaagggtagt agaacgaacc cttctggaac agtttgcaga 2460
taaaaatctt agctatgatg aaagatcaat cagcattatg aaggtggctc aagcgaaact 2520
gaaggaaatt ggtccagatg acatgaatat ggaagagtac aagaagtggc atgaagatta 2580
tagtttgttc cgaaaagtgt ctgtgtatct cctaacaggc ctagaactct atcaaaaagg 2640
aaagtgccaa gaggcacttt cctacctggt atatgcctac cagagcaatg ctgccctgct 2700
gatgaagggg ccccgccggg gggtcaaaga atccgtgatt gctttatacc gaagaaaatg 2760
ccttctggag ctgaatgcca aagcagcttc tctttttgaa acaaatgatg atcactccgt 2820
aactgagggc attaatgtga tgaatgaact gatcatcccc tgcattcacc ttatcattaa 2880
taatgacatt tccaaggatg atctggatgc cattgaggtc atgagaaacc attggtgctc 2940
ttaccttggg caagatattg cagaaaatct gcagctgtgc ctaggggagt ttctacccag 3000
acttctagat ccttctgcag aaatcatcgt cttgaaagag cctccaacta ttcgacccaa 3060
ttctccctat gacctatgta gccgatttgc agctgtcatg gagtcaattc agggagtttc 3120
aactgtgaca gtgaaataag ctcccacatg ttcaaggccc attctggttc ctggctgcct 3180
gcctcttgca cagaagttcg ttgtcatagt gctcaccttg ggaaaaggat taggtgggca 3240
cataagattc cgatcagacc ccaaccatgc tgcatgtgta aagaaggatt gaaaataaaa 3300
ttgcactttt taggtacaaa atcataaaag ctgtttcact agaaaaggca gaaagcagtg 3360
tattaaggtg ttgaattacg ccagaagacc tgaaatgcct tgtacctaca acaatgctta 3420
ggcttttcta agcctcttgc cacttttaaa attatccttc aggcataaat atttttgaca 3480
gcagaataga agaatgattc atgagaacct gaaccagatg aacagctact agttatttta 3540
tcaaatacag atgacattta aaaattctta actacaagag attagaaata taaaccttgc 3600
ctggctcttg ccaggagata acaaaatggg ttgctgatga actgcaccct tttacatgtg 3660
ggtagaatat aagctcacat ggcagtgaga tgttgaaaag tcaaaagaga cctgtctctc 3720
tcctttcttt tctatcttta aaccagaaaa cctcatactc agtcctcagt gaaagaaagt 3780
aaagtattaa ggactttaga cagaagagca ttgtgtaact tgactgaaga tcatccatta 3840
atagttatta ggcatttagg taaaattttc taatacctaa aaattgtcaa aaacagtcaa 3900
tagggctact gctggcccaa agaccattta ggtccacctc ctctttttg ctcttttttt 3960

ttttctgtga cagtttcact gtgtcgccca ggctggcgtt cagtggtgca atctcagctc 4020
actgcaaact ctgtctcctg ggctcaagtg attctcgtgc ctcagcctcc cgaatagctg 4080
gaattacggg catgcaccac cacacctggc taatttttgt atttttaata gagatggggt 4140
ttcaccatat tggccaggct gatctctaac tcctggcctc aagtgatcta tctgcctccc 4200
tcagcctccc aaagtctggg attgcagaca agtcatcgta cccggccttc tttttgccc 4260
ttaaaagtaa gggatgtggg tttgtacaaa aaaaaaaaa aaaaaaaag aggggcggcc 4320
gcgcgattat tgagtctctt gcaacccgcg aatttatttc cgaaccggtt acctgagggc 4380
gttcccagtt tcctaatggt gagtcgtttt acagcttgta gtaatcatga acaagctgtc 4440
ctgtgtgaat tgtttcgttc catccacata tcacacacac aacacggacg gaagacg 4497

<210> 67
<211> 1247
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505933CB1

<400> 67
cggctcgagc cgtcccggag acccggtcgg gagggaggaa ggtggcaaga tggtgttgga 60
aagcactatg gtgtgtgtgg acaacagtga gtatatgcgg aatggagact tcttacccac 120
caggctgcag gcccagcagg atgctgtcaa catagtttgt cattcaaaga cccgcagcaa 180
ccctgagaac aacgtgggcc ttatcacact ggctaatgac tgtgaagtgc tgaccacact 240
caccccagac actggccgta tcctgtccaa gctacatact gtccaaccca agggcaagat 300
caccttctgc acgggcatcc gcgtggccca tctggctctg aagcaccgac aaggcaagaa 360
tcacaagatg cgcatcattg cctttgtggg aagcccagtg gaggacaatg agaaggatga 420
ggtgaacaca gaaaagctga cagcctttgt aaacacgttg aatggcaaag atggaaccgg 480
ttctcatctg gtgacagtgc ctcctgggcc cagtttggct gatgctctca tcagttctcc 540
gattttggct ggtgaaggtg gtgccatgct gggtcttggt gccagtgact ttgaatttgg 600
agtagatccc agtgctgatc ctgagctggc cttggccctt cgtgtatcta tggaagagca 660
gcggcagcgg caggaggagg aggcccggcg ggcagctgca gcttctgctg ctgaggccgg 720
gattgctacg actgggactg aagactcaga cgatgccctg ctgaagatga ccatcagcca 780
gcaagagttt ggccgcactg ggcttcctga cctaagcagt atgactgagg aagagcagat 840
tgcttatgcc atgcagatgt ccctgcaggg agcagagttt ggccaggcgg aatcagcaga 900
cattgatgcc agctcagcta tggacacatc cgagccagcc aaggaggagg atgattacga 960
cgtgatgcag gaccccgagt tccttcagag tgtcctagag aacctcccag gtgtggatcc 1020
caacaatgaa gccattcgaa atgctatggg ctccctggcc tcccaggcca ccaaggacgg 1080
caagaaggac aagaaggagg aagacaagaa gtgagactgg agggaaaggg tagctgagtc 1140
tgcttagggg actgcatggg aagcacggaa tatagggtta gatgtgtgtt atctgtaacc 1200
attacagcct aaataaagct tggcaacttt ttttcaaaaa aaaaaaa 1247

<210> 68
<211> 714
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507064CB1

<400> 68
atgaggggcc ttgtggtatt ccttgcagtc tttgctctct ctgaggtcaa tgccatcacc 60
aggtctccaa cattgtggac cccccaccaga ctgtgggtct gagcacccag gaacctggcg 120
acgtcttcac ctactccgag tttgatggga tcctggggct ggcctatccc tctcttgcct 180
ctgagtagtc agtgccagtg tttgacaaca cgatgcagag gcacctggtg gcccaagacc 240
tgttctcagt ctacatgagt aggaatgacc aggggagcat gctcacgctg agggccattg 300
atctgtcgta ctacacaggc tccctgcact ggatacccat gactgcaaga atactggcag 360
ttcactgtgg acagtttgac atcgactgcg ggcgcctgag cagcattccc acggctgtct 420
tcgagatcca cggcaagaag taccccctgc caccctccgc ctataccagc caggaccagg 480
gcttctgcac cagtggtttc cagggtgact atagttccca gcagtggatc ctggggaatg 540
tcttcatctg ggagtattac agtgtctttg acaggaccaa taaccgtgtg gggctggcga 600
aggctgtctg attgcatcac tggccatgga cctcaatgtg accaaacaca cacacgcaca 660

tagatgagat gtgcaggcag atggttccca ataaacaccg catttctcca aaaa 714

<210> 69
<211> 1008
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1439986CB1

<400> 69
agctagccag ccatgggcca ggattattac tctgtgctcg ggatcactcg caattcagag 60
gatgcccaga tcaagcaggc gtaccgcaga ctcgccctta agcaccaccc gttgaagtca 120
aatgagccgt cttcagcaga gattttcagg caaatagcag aggcctacga cgtgctgagt 180
gaccccatga agagaggcat ctacgacaag tttggagaag agggcctgaa gggtgggatt 240
cctttggagt ttggatccca gaccccatgg acaactggtt acgtcttcca cggcaaacct 300
gaaaaggtgt tccacgagtt ctttggtgga aacaacccct tcagtgagtt ttttgatgca 360
gaaggaagtg aggtagattt gaactttggg gggctccagg gccgaggggt caagaagcag 420
gacccccaag tcgaacggga tctctacctg tccctggagg acttattctt tggctgcacc 480
aaaaaaatta agatctccag aagggtgctg aacgaggatg ggtactcctc caccatcaag 540
gacaagatcc tgaccattga tgtgaagccc ggttggaggc agggcacacg catcaccttt 600
gagaaggaag gggaccaggg ccccaacatc atcccagcag acatcatttt catcgtaaag 660
gagaagctac accctcgctt ccgcagggag aatgacaacc tcttcttcgt gaaccccatc 720
cctcttggca aggctctcac ctgctgcact gtggaggtga ggaccctaga tgaccgtctg 780
ctcaacatcc ccatcaatga catcatccac cccaaatact tcaagaaggt gccaggggag 840
gggatgccat tgccggagga ccccactaag aaaggggatc tcttcatctt cttcgacatc 900
cagttcccca cccgcctcac accccagaag aagcagatgc tgcgccaggc attgctgaca 960
tgactgtggt gggctggagc aggggtgaga ggaggctagc cgggcctc 1008

<210> 70
<211> 2425
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2008979CB1

<400> 70
ccgggccccg accccgtccc gggccccagc gccggccgcc cgcccggtcg ggcgatgaag 60
tgtcactatg aggcgctggg ggtgcggcgc gacgccagcg aggaggagct caagaaggcc 120
tatcggaagc tggccctgaa atggcacccg gataaaaatc tggataatgc cgcagaagca 180
gctgaacaat ttaaattaat ccaagcagca tatgatgtgt tgagtgaccc tcaggaaaga 240
gcatggtatg ataatcatag agaggcccta cttaaaggtg ggtttgatgg cgaatatcaa 300
gatgacagct tagatttgct acgctatttc accgttacct gttattctgg ttatggagat 360
gatgaaaagg gattttacac ggtgtatcgt aatgtttttg aaatgattgc caaggaagaa 420
ctagaatctg tgttagagga agaggttgat gatttcccaa cttttggaga ctcccagagt 480
gactatgata cggtagtcca tcctttctac gcttattggc agagtttctg cactcaaaag 540
aattttgcat ggaaggaaga atatgataca cgacaggctt caaaccgctg ggaaaaacga 600
gccatggaaa aagaaaacaa aaagattcgg gacaaagcaa ggaaagagaa gaatgagctt 660
gtccgtcagc tggtagcttt cattcgtaaa agagataaaa gagtgcaggc gcatcgaaaa 720
cttgtggaag aacagaatgc agagaaggcg aggaaagccg aagagatgag gcggcagcag 780
aagctaaagc aggccaaact ggtggagcag tacagagaac agagctggat gactatggcc 840
aatttggaga aagagctcca ggagatggag gcacggtacg agaaggagtt tggagatgga 900
tcggatgaaa atgaaatgga agaacatgaa ctcaaagatg aggaggatgg taaagacagt 960
gatgaggccg aggacgctga gctctatgat gacctttact gcccagcatg tgacaaatcg 1020
ttcaagacag aaaaggccat gaagaatcac gagaagtcaa agaagcatcg ggaaatggtg 1080
gccttgctaa aacaacagct ggaggaggaa gaagaaaatt tttcaagacc tcaaattgat 1140
gaaaatccat tagatgacaa ttctgaggaa gaaatggaag atgcaccaaa acaaaagctt 1200
tctaaaaaac agaagaaaaa gaaacagaaa ccagcacaga attatgatga caatttcaat 1260
gtaaatggac ctggagaagg agtaaaggtt gatccagaag atactaactt aaatcaagac 1320
agtgccaaag aattggaaga tagtccccag gaaaatgtca gtgtcacaga gatcattaaa 1380

ccatgtgatg atccaaaaag tgaagctaaa agtgttccta aacccaaagg aaagaaaacc 1440
aaagatatga aaaaacctgt cagagtacct gctgaaccac aaacaatgag tgttcttatc 1500
agctgtacaa cctgccatag tgaatttcca tctcggaata aactttttga ccatctaaag 1560
gccacaggtc atgcaagagc accttcatca tcgtctttaa acagcgcaac aagtagtcaa 1620
agcaagaaag agaaacgtaa aaacagatag agattctgcc tgtgcttttg tttgactgtc 1680
tctagatttt gaaaccaaaa aactgaactg aaatcatcta aagagttaaa atttcagtga 1740
tctgcaatta attacattgt ggaagattat tttttatctt gtaaaaacac ttttttggtt 1800
taatatatat ttttaaaaca tttcactagt gattgaattc tacttttgcc atctgaattg 1860
acttgaatgt cttaaaacag gtaaatactg taaagtgtgt attcttgatg tttattggct 1920
catgtggaca gaaatgtaca gggagaatta cattatttta acacacagaa gtgcaacttt 1980
ctgctttatt ttctgaattt cacattactt ttacttaatg cttttgtgtt ttgttaatac 2040
ttcataatat gtgaaaaact cggatctttt aaaaagcatc atagatcatt tttccatatg 2100
acactggttc cgattttaaa aattattttt aaataaccga ttattgatta ctgtattttt 2160
tttctcaaga acagtgatag gtagaaacta attgaacatt tggtagtctt tcaagaatag 2220
tgtctcttca aggttttact tgatttaatt tgatatttta ctggtttacc agtaaggtgt 2280
attgttcagt tttttgctcc gatttgaatt gtggaggtgg aagcaaatta gtttacatgg 2340
catgtcctcc ctaggcacag tgacagctgt aaagtatgac ggaacaaggt agcagatggt 2400
acagaattta tactatttaa gaaga 2425

<210> 71
<211> 856
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 90073157CB1

<400> 71
gaaccactcc ggcccgctag acctgctgct atgtcccggc cgcagcttcg acgctggcgc 60
ctcgtctcta gcccgccgag cggcgtcccg ggtctagcgc tgctggcgct gctggcgctg 120
ctggcgctgc ggctcgcggc cgggaccgac tgcccatgcc cggagcctga gctctgccgc 180
ccgattcgcc accatccaga tttcgaggtc tttgtgtttg atgttggaca gaaaacttgg 240
aaatcttatg attggtcaca gattacaact gtggcaacat ttggaaaata tgactcagaa 300
cttatgtgct acgctcattc aaaaggagcc agagtagtac ttaaaggtaa cctttgatgt 360
agcttggtct ccaaagaaca tagacagaag atgctataat tatactggaa tcgcagatgc 420
ttgtgacttc ctctttgtga tgtcttatga tgaacaaagt cagatctggt cagaatgtat 480
tgcagcagcc aatgctccct ataatcagac attaactgga tataatgact acatcaagat 540
gagcattaat cctaagaaac ttgtaatggg tgttccttgg tatggttatg attatacctg 600
cctgaatctg tctgaggatc atgtttgtac cattgcaagt atggtatgat aaccctcaga 660
gtatttcttt aaaggcaaca tatatacaaa actatcgctt acggggcatt ggcatgtgga 720
atgcaaactg tcttgactac tctggagatg ctgtagccaa acagcaaact gaagaaatgt 780
gggaagtctt aaagccaaag ctgttacaga gatgaacatc ttttgtcaaa ccattaagag 840
ttagaaagat gatctg 856

<210> 72
<211> 1318
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506782CB1

<400> 72
ggaggcttga aactgttaca atatggcttt cctgggactc ttctctttgc tggttctgca 60
aagtatggct acaggggcca ctttccctga ggaagccatt gctgacttgt cagtgaatat 120
gtataatcgt cttagagcca ctggtgaaga tgaaaatatt ctcttctctc cattgagtat 180
tgctcttgca atgggaatga tggaacttgg ggcccaagga tctacccaga aagaaatccg 240
ccactcaatg ggatatgaca gcctaaaaaa tggtgaagaa ttttctttct tgaaggagtt 300
ttcaaacatg gtaactgcta aagagagcca atatgtgatg aaaattgcca attccttgtt 360
tgtgcaaaat ggatttcatg tcaatgagga gtttttgcaa atgatgaaaa aatattttaa 420
tgcagcagta aatcatgtgg acttcagtca aaatgtagcc gtggccaact acatcaataa 480

```
gtgggtggag aataacacaa acagggaatt tagtgatggc tccaatgaag ctggtggtat 540
ctaccaagtc ctagaaatac catatgaagg agatgaaata agcatgatgc tggtgctgtc 600
cagacaggaa gttcctcttg ctactctgga gccattagtc aaagcacagc tggttgaaga 660
atgggcaaac tctgtgaaga agcaaaaagt agaagtatac ctgcccaggt tcacagtgga 720
acaggaaatt gatttaaaag atgttttgaa ggctcttgga ataactgaaa ttttcatcaa 780
agatgcaaat ttgacaggcc tctctgataa taaggagatt tttctttcca aagcaattca 840
caagtccttc ctagaggtta atgaagaagg ctcagaagct gctgctgtct caggaatgat 900
tgcaattagt aggatggctg tgctgtatcc tcaagttatt gtcgaccatc cattttctt 960
tcttatcaga aacaggagaa ctggtacaat tctattcatg ggacgagtca tgcatcctga 1020
aacaatgaac acaagtggac atgatttcga agaactttaa gttactttat ttgaataaca 1080
aggaaaacag taactaagca cattatgttt gcaactggta tatatttagg atttgtgttt 1140
tacagtatat cttaagataa tatttaaaat agttccagat aaaaacaata tatgtaaatt 1200
ataagtaact tgtcaaggaa tgttatcagt attaagctaa tggtcctgtt atgtcattgt 1260
gtttgtgtgc tgttgtttaa ataaaagtac ctattgaaaa aaaaaaaaaa aaaaaaa 1318
```

<210> 73
<211> 2251
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506941CB1

<400> 73

```
cctggagtca gcttaaaaag ctgcttgccg ataccagaaa atatcatggc tacatgatgg 60
ctaaggcacc acatgatttc atgtttgtga agaggaatga tccagatgga cctcattcag 120
acagaatcta ttaccttgcc atgtctggtg agaacagaga aaatacactg ttttattctg 180
aaattcccaa aactatcaat agagcagcag tcttaatgct ctcttggaag cctcttttgg 240
atctttttca ggcaacactg gactatggaa tgtattctcg agaagaagaa ctattaagag 300
aaagaaaacg cattggaaca gtcggaattg cttcttacga ttatcaccaa ggaagtggaa 360
catttctgtt tcaagccggt agtggaattt atcacgtaaa agatggaggg ccacaaggat 420
ttacgcaaca acctttaagg cccaatctag tggaaactag ttgtcccaac atacggatgg 480
atccaaaatt atgccctgct gatccagact ggattgcttt tatacatagc aacgatattt 540
ggatatctaa catcgtaacc agagaagaaa ggagactcac ttatgtgcac aatgagctag 600
ccaacatgga agaagatgcc agatcagctg gagtcgctac ctttgttctc caagaagaat 660
ttgatagata ttctggctat tggtggtgtc caaaagctga aacaactccc agtggtggta 720
aaattcttag aattctatat gaagaaaatg atgaatctga ggtggaaatt attcatgtta 780
catcccctat gttggaaaca aggcagggca gattcattcc gttatcctaa aacaggtaca 840
gcaaatccta aagtcacttt taagatgtca gaaataatga ttgatgctga aggaaggatc 900
catgacatct ttcatgtttt tccccaaagt cacgaagagg aaattgagtt tatttttgcc 960
tctgaatgca aaacaggttt ccgtcattta tacaaaatta catctatttt aaaggaaagc 1020
aaatataaac gatccagtgg tgggctgcct gctccaagtg atttcaagtg tcctatcaaa 1080
gaggagatag caattaccag tggtgaatgg gaagttcttg gccggcatgg atctaatatc 1140
caagttgatg aagtcagaag gctggtatat tttgaaggca ccaaagactc ccctttagag 1200
catcacctgt acgtagtcag ttacgtaaat cctggagagg tgacaaggct gactgaccgt 1260
ggctactcac attcttgctg catcagtcag cactgtgact tctttataag taagtatagt 1320
aaccagaaga atccacactg tgtgtccctt tacaagctat caagtcctga agatgaccca 1380
acttgcaaaa caaaggaatt ttgggccacc attttggatt cagcaggtcc tcttcctgac 1440
tatactcctc cagaaatttt ctcttttgaa agtactactg gatttacatt gtatgggatg 1500
ctctacaagc ctcatgatct acagcctgga aagaaatatc ctactgtgct gttcatatat 1560
ggtggtcctc aggtgcagtt ggtgaataat cggtttaaag gagtcaagta tttccgcttg 1620
aataccctag cctctctagg ttatgtggtt gtagtgatag acaacagggg atcctgtcac 1680
cgagggctta aatttgaagg cgcctttaaa tataaaatgg ttgctattgc tggggcccca 1740
gtcactctgt ggatcttcta tgatacagga tacacggaac gttatatggg tcaccctgac 1800
cagaatgaac agggctatta cttaggatct gtggccatgc aagcagaaaa gttcccctct 1860
gaaccaaatc gtttactgct cttacatggt ttcctggatg agaatgtcca ttttgcacat 1920
accagtatat tactgagttt tttagtgagg gctggaaagc catatgattt acagatctat 1980
cctcaggaga gacacagcat aagagttcct gaatcgggag aacattatga actgcatctt 2040
ttgcactacc ttcaagaaaa ccttggatca cgtattgctg ctctaaaagt gatataattt 2100
tgacctgtgt agaactctct ggtatacact ggctatttaa ccaaatgagg aggtttaatc 2160
aacagaaaac acagaattga tcatcacatt ttgatacctg ccatgtaaca tctactcctg 2220
aaaataaatg tggtgccatg aaaaaaaaaa a 2251
```

<210> 74
<211> 1147
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507072CB1

<400> 74
atgcagggca cccctggagg cgggacgcgc cctgggccat cccccgtgga caggcggaca 60
ctcctggtct tcagctttat cctggcagca gctttgggcc aaatgaattt cacaggggac 120
caggttcttc gagtcctggc caaagatgag aagcagcttt cacttctcgg ggatctggag 180
ggcctgaaac ccagaaggt ggacttctgg cgtggcccag ccaggcccag cctccctgtg 240
gatatgagag ttcctttctc tgaactgaaa gacatcaaag cttatctgga gtctcatgga 300
cttgcttaca gcatcatgat aaaggacatc caggtgctgc tggatgagga aagacgggcc 360
atggcgaaat cccgccggct ggagcgcagc accaacagct tcagttactc atcataccac 420
accctggagg agatatatag ctggattgac aactttgtaa tggagcattc cgatattgtc 480
tcaaaaattc agattggcaa cagctttgaa aaccagtcca ttcttgtcct gaaggaaatg 540
gttctaacag caacccctgc tcagaaactt atcacgggcc ctcccctcag tcggagccgg 600
aggtggctgc catagtgaac ttcatcacag cccatggcaa cttcaaggct ctgatctcca 660
tccacagcta ctctcagatg cttatgtacc cttacggccg attgctggag cccgtttcaa 720
atcagaggga gttgtacgat cttgccaagg atgcggtgga ggccttgtat aaggtccatg 780
ggatcgagta catttttggc agcatcagca ccaccctcta tgtggccagt gggatcaccg 840
tcgactgggc ctacgacagt ggcatcaagt acgccttcag ctttgagctc cgggacactg 900
ggcagtatgg cttcctgctg ccggccacac agatcatccc cacggcccag gagacgtgga 960
tggcgcttcg gaccatcatg gagcacaccc tgaatcaccc ctactagcag cacgactgag 1020
ggcaggaggc tccatccttc tccccaaggt ctgtggctcc tcccgaaacc caagttatgc 1080
atccccatcc ccatgccctc atcccgacct cttagaaaat aaatacaagt ttgaacaggc 1140
gcaaaaa 1147

<210> 75
<211> 800
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507083CB1

<400> 75
ggagaagcca tggctgatga gaaaccatcc aacggtgttc tggtccacat ggtgaagttg 60
ctgatcaaga cctttctaga tggcattttt gatgatttga tggaaaataa tgtgttaaat 120
acagatgaga tacaccttat aggaaaatgt ctaaagtttg tggtgagcaa tgctgaaaac 180
ctggttgatg atatcactga gacagctcaa attgcaggca aaatatttag ggaacacctg 240
tggaattcca aaaaacagct gagttcagat gtatccagtg atggagaaag agaggcgaac 300
atgcctggcc tcaacatccg caacaaagaa ttcaactatc ttcataatcg aaatggttct 360
gaacttgacc ttttggggat gcgagatcta cttgaaaacc ttggatactc agtggttata 420
aaagagaatc tcacagctca gatggtgctg ggattgtttg gttcaccact gacagtggaa 480
aagccagtgc agatactcat ggtcggctct tgcaaggtaa catctgtaat gatgctgtta 540
caaaggctca tgtggaaaag gacttcattg ctttcaaatc ttccacacca cataatgttt 600
cttggagaca tgaaacaaat ggctccgtct tcatttccca aattatctac tacttcagag 660
agtattcttg gagtcatcat ctagaggaaa tttttcaaaa ggttcaacat tcatttgaga 720
ccccaaatat actgacccag ctgcccacca ttgaaagact atccatgaca cgatatttct 780
atctctttcc tgggaattaa 800

<210> 76
<211> 1296
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7509097CB1

<400> 76
ctttaccccta catgcctgag aagccagggc caagatggat cttctcctcg acatcagcta 60
agcctggagg actcttcccc tcagagacca tggagaggga cagccacggg aatgcatctc 120
cagcaagaac accttcagct ggagcatctc cagcccaggc atctccagct gggacacctc 180
caggccgggc atctccagcc caggcatctc cagcccaggc atctccagct gggacacctc 240
cgggccgggc atctccagcc caggcatctc cagctggtac acctccaggc cgggcatctc 300
caggccgggc atctccagcc caggcatctc cagcccaggc atctccagcc cgggcatctc 360
cggctctggc atcactttcc aggtcctcat ccggcaggtc atcatccgcc aggtcagcct 420
cggtgacaac ctccccaacc agagtgtacc ttgttagagc aacaccagtg gggctgtac 480
ccatccgatc atctcctgcc aggtcagcac cagcaaccag ggccaccagg gagagcccag 540
gtacgagcct gcccaagttc acctggcggg agggccagaa gcagctaccg ctcatcgggt 600
gcgtgctcct cctcattgcc ctggtggttt cgctcatcat cctcttccag ttctggcagg 660
gccacacagg gatcaggtac aaggagcaga gggagagctg tcccaagcac gctgttcgct 720
gtgacggggt ggtggactgc aagctgaaga gtgacgagct gggctgcgtg aggtttgact 780
gggacaagtc tctgcttaaa atctactctg ggtcctccca tcagtggctt cccactgcgg 840
actgagggcc atgaccgggc ggatcgtggg aggggcgctg gcctcggata gcaagtggcc 900
ttggcaagtg agtctgcact tcggcaccac ccacatctgt ggaggcacgc tcattgacgc 960
ccagtgggtg ctcactgccg cccactgctt cttcgtgacc cgggagaagg tcctggaggg 1020
ctggaaggtg tacgcgggca ccagcaacct gcaccagttg cctgaggcag cctccattgc 1080
cgagatcatc atcaacagca attacaccga tgaggaggac gactatgaca tcgccctcgc 1140
acttgtcagc atctgtcaac tcatatccgg gccccaaagc ttctgcaggg cagaagtcaa 1200
agactcttaa agatccttac atggaacact tctgttttat aattagggaa actgaagccc 1260
aagggttata aataagtttg ctccaaatga cacatc 1296

<210> 77
<211> 1407
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509118CB1

<400> 77
ctttaccccta catgcctgag aagccaggac caagatggat cttctcctcg acatcagcta 60
agcctggagg actcttcccc tcagagacca tggagaggga cagccacggg aatgcatctc 120
cagcaagaac accttcagct ggagcatctc cagcccaggc atctccagct gggacacctc 180
caggccgggc atctccagcc caggcatctc cagcccaggc atctccagct gggacacctc 240
cgggccgggc atctccagcc caggcatctc cagctggtac acctccaggc cgggcatctc 300
caggccgggc atctccagcc caggcatctc cagcccaggc atctccagcc cgggcatctc 360
cggctctggc atcactttcc aggtcctcat ccggcaggtc atcatccgcc aggtcagcct 420
cggtgacaac ctccccaacc agagtgtacc ttgttagagc aacaccagtg gggctgtac 480
ccatccgatc atctcctgcc aggtcagcac cagcaaccag ggccaccagg gagagcccag 540
gtacgagcct gcccaagttc acctggcggg agggccagaa gcagctaccg ctcatcgggt 600
gcgtgctcct cctcattgcc ctggtggttt cgctcatcat cctcttccag ttctggcagg 660
gccacacagg gatcaggtac aaggagcaga gggagagctg tcccaagcac gctgttcgct 720
gtgacggggt ggtggactgc aagctgaaga gtgacgagct gggctgcgtg aggtttgact 780
gggacaagtc tctgcttaaa atctactctg ggtcctccca tcagtggctt cccatctgta 840
gcagcaactg gaatgactcc tactcagaga agacctgcca gcagctgggt ttcgagagtg 900
ctcaccggac aaccgaggtt gcccacaggg attttgccaa cagcttctca atcttgagat 960
acaactccac catccaggaa agcctccaca ggtctgaatg cccttcccag cggtatatct 1020
ccctccagtg ttcccactgc ggactgaggg ccatgaccgg gcggatcgtg ggaggggcgc 1080
tggcctcgga tagcaagtgg ccttggcaag tgagtctgca cttcggcacc acccacatct 1140
gtggaggcac gctcattgac gcccagtggg tgctcactgc cgcccactgc ttcttcgtga 1200
cccgggagaa ggtcctggag ggctggaagg tgtacgcggg caccagcaac tcatatccgg 1260
gccccaaagc ttctgcaggg cagaagtcaa agactcttaa agatccttac atggaacact 1320
tctgttttat aattagggaa actgaagccc aagggttata aataagtttg ctccaaatga 1380
cacatctcac attacaaatt gatgacg 1407

<210> 78
<211> 1448

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509312CB1

<400> 78
ctcctcgaca tcagctaagc ctggaggact cttcccctca gagaccatgg agagggacag 60
ccacgggaat gcatctccag caagaacacc ttcagctgga gcatctccag cccaggcatc 120
tccagctggg acacctccag gccgggcatc tccagcccag gcatctccag cccaggcatc 180
tccagcccgg gcatctccgg ctctggcatc actttccagg tcctcatccg gcaggtcatc 240
atccgccagg tcagcctcgg tgacaacctc cccaaccaga gtgtaccttg ttagagcaac 300
accagtgggg gctgtaccca tccgatcatc tcctgccagg tcagcaccag caaccagggc 360
caccagggag agcccaggta cgagcctgcc caagttcacc tggcgggagg gccagaagca 420
gctaccgctc atcgggtgcg tgctcctcct cattgccctg gtggtttcgc tcatcatcct 480
cttccagttc tggcagggcc acacagggat caggtacaag gagcagaggg agagctgtcc 540
caagcacgct gttcgctgtg acggggtggt ggactgcaag ctgaagagtg acgagctggg 600
ctgcgtgagg tttgactggg acaagtctct gcttaaaatc tactctgggt cctcccatca 660
gtggcttccc atctgtagca gcaactggaa tgactcctac tcagagaaga cctgccagca 720
gctgggtttc gagagtgctc accggacaac cgaggttgcc cacagggatt ttgccaacag 780
cttctcaatc ttgagataca actccaccat ccaggaaagc ctccacaggt ctgaatgccc 840
ttcccagcgg tatatctccc tccagtgttc ccactgcgga ctgagggcca tgaccgggcg 900
gatcgtggga ggggcgctgg cctcggatag caagtggcct tggcaagtga gtctgcactt 960
cggcaccacc cacatctgtg gaggcacgct cattgacgcc cagtgggtgc tcactgccgc 1020
ccactgcttc ttcgtgaccc gggagaaggt cctggagggc tggaaggtgt acgcgggcac 1080
cagcaacctg caccagttgc ctgaggcagc ctccattgcc gagatcatca tcaacagcaa 1140
ttacaccgat gaggaggacg actatgacat cgccctcatg cggctgtcca agcccctgac 1200
cctgtccggt gagggaatct gcactccccg ctctcctgcc ccccagcccc agcaccctct 1260
gcagccctcg cacttgtcag catctgtcaa ctcatatccg ggccccaaag cttctgcagg 1320
gcagaagtca aagactctta aagatcctta catggaacac ttctgtttta taattaggga 1380
aactgaagcc caagggttat aaataagttt gctccaaatg acacatctca cattacaaat 1440
tgatgacg 1448

<210> 79
<211> 2360
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 90126902CB1

<400> 79
tgccccggag ctcagtccca ccccctccgc cgatgaggcc atgaggcacc gaacggacct 60
gggccaggtc acccccctccc tgtcccatcc ccaatccggg ggactccctg ctgtctggga 120
gctgagggag ccctggggcc ggggtgcagc tgagggcctc ttctagggac ccaactcctg 180
atccagcccc tgccggccgc accttccgca gggtgtcaca gaatcatgga ctgagtccac 240
ttggccccca ctgcacagat ggggaaactg aggcctggga gagtggagtg gctggcctca 300
ggccacacgg aaaggcccca cctcttccag aacctcctgc tcttcctgtg ggccctgctg 360
aactgtggtt tgggggtcag tgctcagggt ccgggcgagt ggacccccgtg ggtgtcctgg 420
acccgctgct ccagctcctg cgggcgtggc gtctccgtgc gcagccggcg ctgcctccgg 480
cttcctgggg aagaaccgtg ctggggagac tcccatgagt accgcctctg ccagttgcca 540
gactgccccc caggggctgt gcccttccga gacctacagt gtgccctgta caatggccgc 600
cctgtcctgg gcacccagaa gacctaccag tgggtgccct tccatggggc gcccaaccag 660
tgcgacctca actgcctggc tgaggggcac gccttctacc acagcttcgg ccgcgtcctg 720
gacggcaccg cctgcagccc gggtgcccag ggggtctgcg tggctggccg ctgccttagc 780
gccggctgtg atgggttgtt gggctcgggt gccctcgagg accgctgtgg ccgctgcgga 840
ggcgccaacg actcgtgcct tttcgtgcag cgcgtgtttc gtgacgccgg tgccttcgct 900
gggtactgga acgtgaccct gatccccgag ggcgccagac acatccgcgt ggaacacagg 960
agccgcaacc acctgggtat cctaggatca ctgatggggg gcgatgggcg ctacgtgctt 1020
aatgggcact gggtggtcag cccaccaggg acctacgagg cggccggcac gcatgtggtc 1080
tacacccgag acacagggcc ccaggagaca ttgcaagcag ccgggcccac ctcccatgac 1140

ctgctcctac aggtcctcct gcaggagccc aaccctggca tcgagtttga gttctggctc 1200
cctcgggagc gctacagccc cttccaggct cgtgtgcagg ccctgggctg gcccctgagg 1260
cagcctcagc cccgggggt ggagcctcag ccccccgcag cccctgctgt caccccctgca 1320
cagaccccaa cgctggcccc agacccctgc ccaccctgcc ctgacacccg cggccgcgcc 1380
caccgactac tccactattg cggcagtgac tttgtgttcc aggcccgagt gctgggccac 1440
caccaccagg cccaggagac ccgctatgag gtgcgcatcc agctcgtcta caagaaccgc 1500
tcgccactgc gggcacgcga gtacgtgtgg gcgccaggcc actgcccctg cccgatgctg 1560
gcaccccacc gggactacct gatggctgtc cagcgtcttg tcagccccga cggcacacag 1620
gaccagctgc tgctgcccca cgccggctac gcccggccct ggagccctgc ggaggacagc 1680
cgcatacgcc tgactgcccg gcgctgtcct ggctgagccc ctgcaggagc cccggccaca 1740
cacagcaaga aagatacatc tgaccagcct caacgtcaac gtatttcccc tctcaccctg 1800
gcttccaggc agctctgaaa tacgtcccac ctgtgcagct atgtgactcc ctcccacaca 1860
cgcttaagac acctctgcat gcagtcaaag ccactgtcac aagccggcag gcactggtga 1920
ggaggcacta aggagactct gacttttatt tcgcctctct ccttggctgc caggaagctc 1980
atagctattt atactcagaa agtttaacgc tgctttcttt ctctttgcgc gcgtcacact 2040
tgcttggaga cactgtcatg aacgagcatg acaccctgct gccctgggta cccagaagat 2100
catctgttta cttcccagac actgtgctgt ctctgctctc tgctactcac acacaccctc 2160
atgtgtgaag ggcagagaca ctgtcacaaa caggcatgcc ccttagaaga catgcctaac 2220
caggcactgt aacgtaccaa cgtaccaatt tccccttttc ccctggctac caggaaactc 2280
ggagacaatc ttttcagcct cagcatttct ggctggattt ccacccatca acacgtgctt 2340
gctcctcctt tttttttttt 2360

<210> 80
<211> 1109
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509352CB1

<400> 80
atggcgtggt ccccaccagc caccctcttt ctgttcctgc tgctgctagg ccagcccccct 60
cccagcaggc cacagtcact gggcaccact aagctccggc tggtgggccc agagagcaag 120
ccagaggagg gccgcctgga ggtgctgcac cagggccagt ggggcaccgt gtgtgatgac 180
aactttgcta tccaggaggc cacagtggct tgccgccagc tgggcttcga agctgccttg 240
acctgggccc acagtgccaa gtacggccaa ggggagggac ccatctggct ggacaatgtg 300
cgctgtgtgg gcacagagag ctccttggac cagtgccaca ggcattacca cagcattgag 360
gtcttcaccc actacgacct cctcactctc aatggctcca aggtggctga ggggcacaag 420
gccagcttct gtctggagga cacaaactgc cccacaggac tgcagcggcg ctacgcatgt 480
gccaactttg gagaacaggg agtgactgta ggctgctggg acacctaccg gcatgacatt 540
gattgccagt gggtggatat cacagatgtg ggcccccggga attatatctt ccaggtgatt 600
gtgaaccccc actatgaagt ggcagagtca gatttctcca acaatatgct gcagtgccgc 660
tgcaagtatg atgggcaccg ggtctggctg cacaactgcc acacagggaa ttcatacccca 720
gccaatgcag aactctccct ggagcaggaa cagcgtctca ggaacaacct catctgaagc 780
tgtcactgca cactcctagc tgctgccgat acaccagata cctcagctta ttggagccat 840
gcccttcaca gagtcccaac tcagaggaaa agggccagtg ccaaggggca ccaagaacct 900
gctcaggaag ccttttgatg gcaagatcac caatccagat ggtattgctc cctcaggatg 960
gctctgggcc tgcccctaag ggcctgtggc ctatggaata tgtcctccag gctttgctca 1020
gctgagctcc tcttctgtaa ggaaacccag tcatccctga atcttgccac agagatccgg 1080
gattcaggag ctctcagttt cttaaggag 1109

<210> 81
<211> 905
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509341CB1

<400> 81
caacaccaag cgcatggact ctgaccccag tgcagtgtct gtagggaaga ggagccatgg 60

ggcttcgggc aggccccatc ctgcttctgc tgctgtggct gctgccaggg gcccattggg 120
atgtgctgcc ttcagaatgc ggccactcca aggaggccgg gaggattgtg ggaggccaag 180
acacccagga aggacgctgg ccgtggcagg ttggcctttg gttgacctca gtggggcatg 240
tatgtggggg ctccctcatc cacccacgct gggtgctcac agccgcccac tgcttcctga 300
ggtctgagga tcccgggctc taccatgtta aagtcggagg gctgacaccc tcactttcag 360
agccccactc ggccttggtg gctgtgagga ggctcctggt ccactcctca taccatggga 420
ccaccaccag cggggacatt gccctgatgg agctggactc cccccttgcag gcctcccagt 480
tcagccccat ctgcctccca ggaccccaga cccccctcgc cattgggacc gtgtgctggg 540
gtgactccgg ggggccgctg gtctgcccca tcaatgatac gtggatccag gccggcattg 600
tgagctgggg attcggctgt gcccggcctt tccggcctgg tgtctacacc caggtgctaa 660
gctacacaga ctggattcag agaaccctgg ctgaatctca ctcaggcatg tctggggccc 720
gcccaggtgc cccaggatcc cactcaggca cctccagatc ccacccagtg ctgctgcttg 780
agctgttgac cgtatgcttg cttgggtccc tgtgaaccat gagccatgga gtccgggatc 840
cccttctgg taggattgat ggaatctaat aataaaaact gtaggttttt tatgtgtatt 900
ttaaa 905

<210> 82
<211> 626
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509367CB1

<400> 82
caacaccaag cgcatggact ctgaccccag tgcagtgtct gtagggaaga ggagccatgg 60
ggcttcgggc aggccccatc ctgcttctgc tgctgtggct gctgccaggg gcccattggg 120
atgtgctgcc ttcagaatgc ggccactcca aggaggccgg gaggattgtg ggaggccaag 180
acacccagga aggacgctgg ccgtggcagg ttggcctgtg gttgacctca gtggggcatg 240
tatgtggggg ctccctcatc cacccacgct gggtgctcac agccgcccac tgcttcctga 300
gggtgactcc ggggggccgc tggtctgccc catcaatgat acgtggatcc aggccggcat 360
tgtgagctgg ggattcggct gtgcccggcc tttccggcct ggtgtctaca cccaggtgct 420
aagctacaca gactggattc agagaaccct ggctgaatct cactcaggca tgtctggggc 480
ccgcccaggt gccccaggat cccactcagg cacctccaga tcccacccag tgctgctgct 540
tgagctgttg accgtatgct tgcttgggtc cctgtgaacc atgagccatg gagtccggga 600
tccccttttct ggtaggattg atggaa 626

<210> 83
<211> 4095
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7500455CB1

<400> 83
cccacgcgtc cggggggtgag ctgggctggc ttccgtcctg gtagccaagg ctaatcctcc 60
ctcgagttct tgggagatgg gcatttggcg agaaggctgg cgttagtgaa gcgcgcccgg 120
cgtcacggaa agaagcactt gtaaggaaat atagcatcca ttgtgaaagt ggaaaagtaa 180
agataattca tcatgcctgc tgtggcttca gttcctaaag aactctacct cagttcttca 240
ctaaaagacc ttaataagaa gacagaagtt aaaccagaga aaataagcac taagaggatt 300
atttccattc aatacttgga cctggaaaca tcaaaaaagc tgtcgaagaa gctgaaagac 360
tctctgaaag ccttaaatta agatatgaag aagctgaagt ccggaaaaaa cttgaggaaa 420
aagacaggca ggaggaagca cagcggctac aacaaaaaag gcaggaaaca ggaagagagg 480
atggtggcac attggctaaa ggctctttgg agaatgtttt ggattccaaa gacaaaaccc 540
aaaagagcaa tggtgaaaag aatgaaaaat gtgagaccaa agagaaagga gcaatcacag 600
caaaggaact atacacaatg atgacggata aaaacatcag cttgattata atggatgctc 660
gaagaatgca ggattatcag gattcctgta ttttacattc tctcagtgtt cctgaagaag 720
ccatcagtcc aggagtcact gctagttgga ttgaagcaca cctgccagat gattctaaag 780
acacatggaa gacgaggggg aatgtggagt atgtggtact tcttgactgg tttagttctg 840
ccaaagattt acagattgga acaactctcc ggagtctgaa agatgcactt ttcaagtggg 900

aaagtaaaac tgtcctgcgc aatgagcctt tggttttaga gggaggctat gaaaactggc 960
tcctttgtta tccccagtat acaacaaatg ctaaggtcac tccaccccca cgacgccaga 1020
atgaagaggt gtctatctca ttggatttta cttatccctc attggaagaa tcaattcctt 1080
ctaaacctgc tgcccagacg ccacctgcat ctatagaagt agatgaaaat atagaattga 1140
taagtggtca aaatgagaga atgggaccac tgaatatatc aactccagtt gaaccagttg 1200
ctgcttctaa atctgatgtt tcacccataa ttcagccagt gcctagtata aagaatgttc 1260
cacagattga tcgtactaaa aaaccagcag tcaaattgcc tgaagagcat agaataaaat 1320
ctgaaagtac aaaccatgag caacaatctc ctcagagtgg aaaagttatt cctgatcgtt 1380
ccaccaagcc agtagttttt tctccaactc tcatgttaac agatgaagaa aaggctcgta 1440
ttcatgcaga aactgctctt ctaatggaaa aaaacaaaca agaaaaagaa cttcgggaaa 1500
ggcagcaaga ggaacagaaa gagaaactga ggaaggaaga acaagaacaa aaagccaaaa 1560
agaaacaaga agctgaagaa aatgaaatta cagagaagca acaaaaagca aaagaagaaa 1620
tggagaagaa agaaagtgaa caggccaaga aagaagataa agaaacctca gcaaagaggg 1680
gcaaagaaat aacaggagta aaaagacaaa gtaaaagtga acatgaaact tctgatgcca 1740
agaaatctgt agaagatagg gggaaaaggt gtccaacccc agaaatacag aaaaagtcaa 1800
caggagatgt gccccataca tctgtgacag gggattcagg ttcaggcaag ccatttaaga 1860
ttaaaggaca accagaaagt ggaattctaa ggacaggaac ttttagagag gatacagacg 1920
ataccgaaag aaataaagct caacgagaac ctttgacaag agcacgaagt gaagaaatgg 1980
ggaggatcgt aacaggactg ccttcaggct gggccaagtt tcttgaccca atcactggaa 2040
cctttcgtta ttatcattca cccaccaaca ctgttcatat gtacccaccg gaaatggctc 2100
cttcatctgc acctccttcc acccctccaa ctcataaagc caagccacag attcctgctg 2160
agcgggatag ggaaccttcc aaactgaagc gctcctactc ctccccagat ataacccagg 2220
ctattcaaga ggaagagaag aggaagccaa cagtaactcc aacagttaat cgggaaaaca 2280
agccaacatg ttatcctaaa gctgagatct caaggctttc tgcttctcag attcggaacc 2340
tcaatcctgt ttttggaggt tctggaccag ctcttactgg acttcgtaac ttaggaaata 2400
cttgttatat gaactcaata ttgcagtgcc tatgtaacgc tccacatttg gctgattatt 2460
tcaaccgaaa ctgttatcag gatgatatta acaggtcaaa tttgttgggg cataaaggtg 2520
aagtggcaga agaatttggt ataatcatga aagccctgtg gacaggacag tatagatata 2580
tcagtccaaa ggactttaaa atcaccattg ggaagatcaa tgaccagttt gcaggataca 2640
gtcagcaaga ttcacaagaa ttgcttctgt tcctaatgga tggtctccat gaagatctaa 2700
ataaagctga taatcggaag agatataaag aagaaaataa tgatcatctc gatgacttta 2760
aagctgcaga acatgcctgg cagaaacaca agcagctcaa tgagtctatt attgttgcac 2820
tttttcaggg tcaattcaaa tctacagtac agtgcctcac atgtcacaaa aagtctagga 2880
catttgaggc cttcatgtat ttgtctctac cactagcatc cacaagtaaa tgtacattac 2940
aggattgcct tagattattt tccaaagaag aaaaactcac agataacaac agattttact 3000
gcagtcattg cagagctcga cgggattctc taaaaaagat agaaatctgg aagttaccac 3060
ctgtgctttt agtgcatctg aaacgttttt cctacgatgg caggtggaaa caaaaattac 3120
agacatctgt ggacttcccg ttagaaaatc ttgacttgtc acagtatgtt attggtccaa 3180
agaacaattt gaagaaatat aatttgtttt ctgtttcaaa tcactacggt gggctggatg 3240
gaggccacta cacagcctat tgtaaaaatg cagcaagaca acggtggttt aagtttgatg 3300
atcatgaagt ttctgatatc tccgtttctt ctgtgaaatc ttcagcagct tatatcctct 3360
tttatacttc attgggacca cgagtaactg atgtagccac ataaggagac ataggttata 3420
aactagttat cttttaaaag gctcagcaac acaactcttg aaatgcttat caggataatg 3480
gtagctatag ctggccattt agaggaattc taggacagtg ggagctgtgt tactagcact 3540
atataattcc ggtcagtgct gacaaataac atttaacaag tattgcagta atcatcactt 3600
acaggtacca tttatttcaa aacaactttt ttagtctgct ccaaagttaa aataattaac 3660
tagctaagca ttattattcg actggtctaa aaactattgt tatctttttt ttttcctttt 3720
cactgttatg gccttttcac atttctaaat cccatcttga tatactatga atactctaga 3780
atgatgtaaa gcagatagga atgtatgtgt acatatttat tgcatacttg cacatcaaat 3840
cgatgtacat agtttaacac gtggtccttt tgtgaaacct agaactcaga ggattgcttt 3900
ttttctttca gcctattttg agttaacttc agtgctttct tagggaaatg acagggcaaa 3960
gcaatttttc tgttggcttt gggctgtatt tgtgcactaa atctttattc taaaaaaaaa 4020
aatggaaact ttaatttttt taaaacgaga atttcattta cagctacatt aaaatcttaa 4080
tgagaaaaat aatta 4095

<210> 84
<211> 1308
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510401CB1

<400> 84
ggcacttccg cttcccctcg gctttcttct cgtcggtgtt cccggctgct atagagccgg 60
gtgagagagc gagcgcccgt cggcgggtgt cgagggcggg ttgcctcgcg ctgacccttc 120
ccgccctcct tctcgtcaca caccaggtcc ccgcggaagc cgcggtgtcg gcgccatggc 180
ggagctgacg gctcttgaga gtctcatcga gatgggcttc cccaggggac gcgcggagaa 240
ggctctggcc ctcacaggga accagggcat cgaggctgcg atggactggc tgatggagca 300
cgaagacgac cccgatgtgg acgagccttt agagactccc cttggacata tcctgggacg 360
ggagcccact tcctcagagc aaggcggcct tgaaggatct ggttctgctg ccggagaagg 420
caaacccgct ttgagtgaag aggaaagaca ggaacaaact aagaggatgt tggagctggt 480
ggcccagaag cagcgggaga gtgaagaaag agaggaacgg gaggcattgg aacgggaacg 540
gcagcgcagg agacaagggc aagagttgtc agcagcacga cagcggctac aggaagatga 600
gatgcgccgg gctgctgagg agaggcggag ggaaaaggcc gaggagttag cagccagaca 660
aagagttaga gaaaagatcg agagggacaa agcagagaga gccaagaagg taggtgactg 720
agaagacgcc acgggtaatg gagtgggcag gttagtaaat cttttccagc ttcagaggga 780
aagcatgggc ttatatgatc tgtttcacct gaggatgggt gagtacctga ttgactgttt 840
tactttgccc tttagtatgg tggcagtgtg ggctctcagc cacccccagt ggcaccagag 900
ccaggtcctg ttccctcttc tcccagccag gagcctccca ccaagcggga gtatgaccag 960
tgtcgcatac aggtcaggct gccagatggg acctcactga cccagacgtt ccgggcccgg 1020
gaacagctgg cagctgtgag gctctatgtg gagctccacc gtggggagga actaggtggg 1080
ggccaggacc ctgtgcaatt gctcagtggc ttccccagac gggccttctc agaagctgac 1140
atggagcggc ctctgcagga gctgggactc gtgccttctg ctgttctcat tgtggccaag 1200
aaatgtccca gctgagggcc tttgtcccat tgtccctctg tgaccccttc atctttgata 1260
aagcactgac atctccttcc taataaatag accctgagtt ctgtaaaa 1308

<210> 85
<211> 1196
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504702CB1

<400> 85
cagcgatgtc tccaccaccg ctgctgcaac ccctgctgct gctgctgcct ctgctgaatg 60
tggagccttc cggggccaca ctgatccgca tccctcttca tcgagtccaa cctggacgca 120
ggatcctgaa cctactgagg ggatggagag aaccagcaga gctcccccaag ttgggggccc 180
catcccctgg ggacaagccc atcttcgtac ctctctcgaa ctacagggat ggttacacca 240
ccgatttgat cccaaagcct ctagctcctt ccaggccaat gggaccaagt ttgccattca 300
atatggaact gggcgggtag atggaatcct gagcgaggac aagctgacta ttggtggaat 360
caagggtgca tcagtgattt tcggggaggc tctctgggag cccagcctgg tcttcgcttt 420
tgcccatttt gatgggatat tgggcctcgg ttttcccatt ctgtctgtgg aaggagttcg 480
gcccccgatg gatgtactgg tggagcaggg gctattggat aagcctgtct tctcctttta 540
cctcaacagg gaccctgaag agcctgatgg aggagagctg gtcctggggg gctcggaccc 600
ggcacactac atcccacccc tcaccttcgt gccagtcacg gtccctgcct actggcagat 660
ccacatggag cgtgtgaagg tgggcccagg gctgactctc tgtgccaagg gctgtgctgc 720
catcctggat acgggcacgt ccctcatcac aggacccact gaggagatcc gggccctgca 780
tgcagccatt gggggaatcc ccttgctggc tggggagtac atcatcctgt gctcggaaat 840
cccaaagctc cccgcagtct ccttccttct tggggggtc tggtttaacc tcacggccca 900
tgattacgtc atccagacta ctcgaaatga cgtcttcttg gggacgtatg tggccgtctt 960
cgaccgcggg gacatgaaga gcagcgcccg ggtgggcctg gcgcgcgctc gcactcgcgg 1020
agcggacctc ggatggggag agactgcgca ggcgcagttc cccgggtgac gcccaagtga 1080
agcgcatgcg cagcgggtgg tcgcggaggt cctgctaccc agtaaaaatc cactatttcc 1140
attgaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 1196

<210> 86
<211> 1419
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509113CB1

<400> 86
ctttgctcca gcagcctgtc ccagtgagga cagggacaca gtactcggcc acaccatggg 60
gcgcccactg cacctcgtcc tgctcagtgc ctccctggct ggcctcctgc tgctcgggga 120
aagtctgttc atccgcaggg agcaggccaa caacatcctg gcgagggtca cgagggccaa 180
ttcctttctt gaagagatga agaaaggaca cctcgaaaga gagtgcatgg aagagacctg 240
ctcatacgaa gaggcccgcg aggtctttga ggacagcgac aagacgaatg aattctggaa 300
taaatacaaa gatggcgacc agtgtgagac cagtccttgc cagaaccagg gcaaatgtaa 360
agacggcctc ggggaataca cctgcacctg tttagaagga ttcgaaggca aaaactgtga 420
attatggccc tacccctgtg ggaaacagac cctggaacgc aggaagaggt cagtggccca 480
ggccaccagc agcagcgggg aggcccctga cagcatcaca tggaagccat atgatgcagc 540
cgacctggac ccccaccgaga acccccttcga cctgcttgac ttcaaccaga cgcagcctga 600
gaggggcgac aacaacctca ccaggatcgt gggaggccag gaatgcaagg acggggagtg 660
tccctggcag gccctgctca tcaatgagga aaacgagggt ttctgtggtg gaactattct 720
gagcgagttc tacatcctaa cggcagccca ctgtctctac caagccaaga gattcaaggt 780
gagggtaggg gaccggaaca cggagcagga ggagggcggt gaggcggtgc acgaggtgga 840
ggtggtcatc aagcacaacc ggttcacaaa ggagacctat gacttcgaca tcgccgtgct 900
ccggctcaag acccccatca ccttccgcat gaacgtggcg cctgcctgcc tccccgagcg 960
tgactgggcc gagtccacgc tgatgacgca gaagacgggg attgtgagcg gcttcgggcg 1020
cacccacgag aagggccggc agtccaccag gctcaagatg ctggaggtgc cctacgtgga 1080
ccgcaacagc tgcaagctgt ccagcagctt catcatcacc cagaacatgt tctgtgccgg 1140
ctacgacacc aagcaggagg atgcctgcca gggggacagc gggggcccgc acgtcacccg 1200
cttcaaggac acctacttcg tgacaggcat cgtcagctgg ggagagggct gtgcccgtaa 1260
ggggaagtac gggatctaca ccaaggtcac cgccttcctc aagtggatcg acaggtccat 1320
gaaaaccagg ggcttgccca aggccaagag ccatgccccg gaggtcataa cgtcctctcc 1380
attaaagtga gatcccactc aaaaaaaaaa aaaagggggg 1419

<210> 87
<211> 1200
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509140CB1

<400> 87
gcagcctgtc ccagtgagga cagggacaca gtactcggcc acaccatggg gcgcccactg 60
cacctcgtcc tgctcagtgc ctccctggct ggcctcctgc tgctcgggga aagtctgttc 120
atccgcaggg agcaggccaa caacatcctg gcgagggtca cgagggccaa ttcctttctt 180
gaagagatga agaaaggaca cctcgaaaga gagtgcatgg aagagacctg ctcatacgaa 240
gaggcccgcg aggtctttga ggacagcgac aagacgaatg aattctggaa taaatacaaa 300
gatggcgacc agtgtgagac cagtccttgc cagaaccagg gcaaatgtaa agacggcctc 360
ggggaataca cctgcacctg tttagaagga ttcgaaggca aaaactgtga attattcaca 420
cggaagctct gcagcctgga caacggggac tgtgaccagt tctgccacga ggaacagaac 480
tctgtggtgt gctcctgcgc ccgcgggtac accctggctg acaacggcaa ggcctgcatt 540
cccacagggc cctacccctg tgggaaacag accctggaac gcaggaagag gtcagtggcc 600
caggccacca gcagcagcgg ggaggcccct gacagcatca catggaagcc atatgatgca 660
gccgacctgg accccaccga gaacccccttc gacctgcttg acttcaacca gacgcagcct 720
gagaggggcg acaacaacct caccaggatc gtgggaggcc aggaatgcaa ggacggggag 780
tgtccctggc aggccctgct catcaatgag gaaaacgaga agggccggca gtccaccagg 840
ctcaagatgc tggaggtgcc ctacgtggac cgcaacagct gcaagctgtc cagcagcttc 900
atcatcaccc agaacatgtt ctgtgccggc tacgacacca agcaggagga tgcctgccag 960
ggggacagcg ggggcccgca cgtcacccgc ttcaaggaca cctacttcgt gacaggcatc 1020
gtcagctggg gagagggctg tgcccgtaag gggaagtacg ggatctacac caaggtcacc 1080
gccttcctca agtggatcga caggtccatg aaaaccaggg gcttgcccaa ggccaagagc 1140
catgccccgg aggtcataac gtcctctcca ttaaagtgag atcccactca aaaaaaaaaa 1200

<210> 88
<211> 1982
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 7509223CB1

<220>
<221> unsure
<222> (1) ... (1982)
<223> a, t, c, g, or other

<400> 88
gttttgggta tcgtaccgtt tgattaccac ccttggagcg catttgaatt atgccatgtg 60
cggaaacgcc tgcagtaccg tccggatccc ggtcgacaga cattccagag aatcctgttt 120
tagaccacct tgcgctgaga gcacacccgt gggtcaaagg cagccacccg gggtcaaag 180
ggcaccatca ggtactcccc agggaaggct ttgcgccacc agtcacctgc aaccccgcct 240
cacctccgat gcctgctgtg cccagggtgt cccgctcata gcgacggcct gtgcgttcat 300
acgacctcga gcgctctcgt cgctgatctc tgtagtaatg acttttcttc ttgtattttc 360
ccggattatc cgcccgctcc cgtgacctgc ttcgagaacg ggaagaactc cggctccgag 420
acttgtgtgg cttctggggg tacttgcagt ccttggactt ccgggagccc cgaatctcag 480
agcctttgta gggagcgctg cggggggcct gtctcggtgg ggagtcactc ctgctccggg 540
agcggctctg cgactgannn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 600
nnngggacac agtactcggc cacaccatgg ggcgcccact gcacctcgtc ctgctcagtg 660
cctccctggc tggcctcctg ctgctcgggg aaagtctgtt catccgcagg gagcaggcca 720
acaacatcct ggcgagggtc acgagggcca attcctttct tgaagagatg aagaaaggac 780
acctcgaaag agagtgcatg gaagagacct gctcatacga agaggcccgc gaggtctttg 840
aggacagcga caagacgaat gaattctgga ataaatacaa agatggcgac cagtgtgaga 900
ccagtccttg ccagaaccag ggcaaatgta aagacggcct cggggaatac acctgcacct 960
gtttagaagg attcgaaggc aaaaactgtg aattattcac acggaagctc tgcagcctgg 1020
acaacgggga ctgtgaccag ttctgccacg aggaacagaa ctctgtggtg tgctcctgcg 1080
cccgcgggta caccctggct gacaacggca aggcctgcat tcccacaggg ccctacccct 1140
gtgggaaaca gaccctggaa cgcaggaaga ggtcagtggc ccaggccacc agcagcagcg 1200
gggaggcccc tgacagcatc acatggaagc catatgatgc agccgacctg gaccccaccg 1260
agaacccctt cgacctgctt gacttcaacc agacgcagcc tgagaggggc gacaacaacc 1320
tcaccaggat cgtgggaggc caggaatgca aggacgggga gtgtccctgg caggccctgc 1380
tcatcaatga ggaaaacgag gtggaggtgg tcatcaagca caaccggttc acaaaggaga 1440
cctatgactt cgacatcgcc gtgctccggc tcaagacccc catcaccttc cgcatgaacg 1500
tggcgcctgc ctgcctcccc gagcgtgact gggccgagtc cacgctgatg acgcagaaga 1560
cggggattgt gagcggcttc gggcgcaccc acgagaaggg ccggcagtcc accaggctca 1620
agatgctgga ggtgccctac gtggaccgca acagctgcaa gctgtccagc agcttcatca 1680
tcacccagaa catgttctgt gccggctacg acaccaagca ggaggatgcc tgccaggggg 1740
acagcggggg cccgcacgtc acccgcttca aggacaccta cttcgtgaca ggcatcgtca 1800
gctggggaga gggctgtgcc cgtaagggga agtacgggat ctacaccaag gtcaccgcct 1860
tcctcaagtg gatcgacagg tccatgaaaa ccaggggctt gcccaaggcc aagagccatg 1920
ccccggaggt cataacgtcc tctccattaa agtgagatcc cactcaaaaa aaaaaaaaag 1980
gg 1982

<210> 89
<211> 1282
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509272CB1

<400> 89
gccagggacc aagatggatc ttctcctcga catcagctaa gcctggagga ctcttcccct 60
cagagaccat ggagagggac agccacggga atgcatctcc agcaagaaca ccttcagctg 120
gagcatctcc agcccaggca tctccagctg ggacacctcc aggccgggca tctccagccc 180
aggcatctcc agcccaggca tctccagctg ggacacctcc gggccgggca tctccagccc 240
aggcatctcc agctggtaca cctccaggcc gggcatctcc aggccgggca tctccagccc 300
aggcatctcc agcccaggca tctccagccc gggcatctcc ggctctggca tcactttcca 360
ggtcagcacc agcaaccagg gccaccaggg agagcccagg tacgagcctg cccaagttca 420
cctggcggga gggccagaag cagctaccgc tcatcgggtg cgtgctcctc ctcattgccc 480
tggtggtttc gctcatcatc ctcttccagt tctggcaggg ccacacaggg atcaggtaca 540

aggagcagag ggagagctgt cccaagcacg ctgttcgctg tgacggggtg gtggactgca 600
agctgaagag tgacgagctg ggctgcgtga ggtttgactg ggacaagtct ctgcttaaaa 660
tctactctgg gtcctcccat cagtggcttc ccatctgtag cagcaactgg aatgactcct 720
actcagagaa gacctgccag cagctgggtt tcgagagtgc tcaccggaca accgaggttg 780
cccacaggga ttttgccaac agcttctcaa tcttgagata caactccacc atccaggaaa 840
gcctccacag gtctgaatgc ccttcccagc ggtatatctc cctccagtgt tcccactgcg 900
gactgagggc catgaccggg cggatcgtgg gaggggcgct ggcctcggat agcaagtggc 960
cttggcaagt gagtctgcac ttcggcacca cccacatctg tggaggcacg ctcattgacg 1020
cccagtgggt gctcactgcc gcccactgct tcttcgtgac ccgggagaag gtcctggagg 1080
gctggaaggt gtacgcgggc accagcaacc tgcaccagtt gcctgaggca gcctccattg 1140
ccgagatcat catcaacagc aattacaccg atgaggagga agtcaaagac tcttaaagat 1200
ccttacatgg aacacttctg ttttataatt agggaaactg aagcccaagg gttataaata 1260
agtttgctcc aaatgacaca tc 1282

<210> 90
<211> 1337
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509327CB1

<400> 90
gccagggacc aagatggatc ttctcctcga catcagctaa gcctggagga ctcttcccct 60
cagagaccat ggagagggac agccacggga atgcatctcc agcaagaaca ccttcagctg 120
gagcatctcc agcccaggca tctccagctg ggacacctcc aggccgggca tctccagccc 180
aggcatctcc agcccaggca tctccagctg ggacacctcc gggccgggca tctccagccc 240
aggcatctcc agctggtaca cctccaggcc gggcatctcc aggccgggca tctccagccc 300
aggcatctcc agcccaggca tctccagccc gggcatctcc ggctctggca tcactttcca 360
ggtcctcatc cggcaggtca tcatccgcca ggtcagcctc ggtgacaacc tccccaacca 420
gagtgtacct tgttagagca acaccagtgg gggctgtacc catccgatca tctcctgcca 480
ggtcagcacc agcaaccagg gccaccaggg agagcccagg tacgagcctg cccaagttca 540
cctggcggga gggccagaag cagctaccgc tcatcgggtg cgtgctcctc ctcattgccc 600
tggtggtttc gctcatcatc ctcttgctca ccggacaacc gaggttgccc acagggattt 660
tgccaacagc ttctcaatct tgagatacaa ctccaccatc caggaaagcc tccacaggtc 720
tgaatgccct tcccagcggt atatctccct ccagtgttcc cactgcggac tgagggccat 780
gaccgggcgg atcgtgggag gggcgctggc ctcggatagc aagtggcctt ggcaagtgag 840
tctgcacttc ggcaccaccc acatctgtgg aggcacgctc attgacgccc agtgggtgct 900
cactgccgcc cactgcttct tcgtgacccg ggagaaggtc ctggagggct ggaaggtgta 960
cgcgggcacc agcaacctgc accagttgcc tgaggcagcc tccattgccg agatcatcat 1020
caacagcaat tacaccgatg aggaggacga ctatgacatc gccctcatgc ggctgtccaa 1080
gcccctgacc ctgtccggtg agggaatctg cactccccgc tctcctgccc cccagcccca 1140
gcaccctctg cagccctcgc acttgtcagc atctgtcaac tcatatccgg gccccaaagc 1200
ttctgcaggg cagaagtcaa agactcttaa agatccttac atggaacact tctgttttat 1260
aattagggaa actgaagccc aagggttata aataagtttg ctccaaatga cacatctcac 1320
attacaaatt gatgacg 1337

<210> 91
<211> 843
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504677CB1

<400> 91
cgcaggacaa cgcctgccac gcgtcggaac tcggccgcgg gacatccacg gggcgcgagt 60
gacacgcggg agggagagca gtgttctgct ggagccgatg ccaaaaacca tgcatttctt 120
attcagattc attgttttct tttatctgtg gggccttttt actgctcaga gacaaaagaa 180
agaggagagc accgaagaag tgaaaataga agttttgcat cgtccagaaa actgctctaa 240
gacaagcaag aagggagacc tactaaatgc ccattatgac ggctacctgg ctaaagacgg 300

ctcgaaattc tactgcagca gaaggcaaga ttccaccgga tgctacattg atttttgaga 360
ttgaacttta tgctgtgacc aaaggaccac ggagcattga gacatttaaa caaatagaca 420
tggacaatga caggcagctc tctaaagccg agataaacct ctacttgcaa agggaatttg 480
aaaaagatga gaagccacgt gacaagtcat atcaggatgc agttttagaa gatattttta 540
agaagaatga ccatgatggt gatggcttca tttctcccaa ggaatacaat gtataccaac 600
acgatgaact atagcatatt tgtatttcta cttttttttt tagctattta ctgtacttta 660
tgtataaaac aaagtcactt ttctccaagt tgtatttgct atttttcccc tatgagaaga 720
tattttgatc tccccaatac attgattttg gtataataaa tgtgaggctg ttttgcaaac 780
ttaacttgca ggaatggtat gactgtgttt cctactgctt tattctgtaa acaagaattg 840
tag 843

<210> 92
<211> 3360
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504534CB1

<400> 92
ggggcggggc cgcggaggag atggcggaag gtggaggctg ccgtgagcga ccggatgcgg 60
agactcagaa gtccgagctt ggacccttaa tgaggaccac actccaacgc ggggcgcagt 120
ggtatcttat tgacagccgg tggttcaagc agtggaagaa gtatgtgggc tttgacagct 180
ggggcatgta caatgtgggt gaacataacc tatttcctgg cccaatagac aactctgggc 240
tattttcaga tcctgagagt cagaccttga aagaacactt aattgatgaa ttggactatg 300
tattggtccc taccgaggcg tggaataaac tactaaactg gtacggctgt gtagaaggcc 360
agcaacccat cgtcagaaaa gttgtggagc atggcctgtt tgtcaagcac tgcaaagtcg 420
aggtgtattt gctggaactg aagctctgtg agaacagtga ccccaccaat gtgctgagtt 480
gccatttcag caaggcagac accattgcaa ccatcgagaa agagatgcgg aagctattca 540
acatccctgc ggagcgtgaa acacggctct ggaacaaata catgagcaac acctacgagc 600
agttgagcaa gctagacaac actgtccagg atgctgggct ataccagggt caggtgctag 660
taattgagcc tcaaaatgaa gatggcacat ggcccaggca gaccttgcag tcaaaatcaa 720
gcactgcgcc tagcagaaat tttactacct ctccaaaatc atcagcaagt ccctattcct 780
cagtgtctgc ctctctcatt gcaaatggtg atagcactag cacctgtggg atgcacagtt 840
ccggtgtcag caggggtgga tctggctttt ctgcttcgta taattgtcag gagccaccat 900
cctctcatat acaacctggg ctctgtggac ttggaaacct gggaaacacc tgcttcatga 960
actccgcttt gcagtgtttg agcaacactg caccactgac tgactacttt ctcaaagatg 1020
agtatgaagc cgaaatcaac agagacaacc ctctggggat gaaaggggaa attgcagaag 1080
cctatgctga actcattaag cagatgtggt ctggaaggga tgcccatgtg gcacctcgca 1140
tgttcaaaac tcaagtagga cgttttgctc ctcaattttc tggctaccag caacaagatt 1200
ctcaggagct gctggccttt cttctagatg gattgcatga agatctgaac cgggtaaaga 1260
aaaagcccta cttggagctg aaggatgcca atgggcggcc agatgcggtg gtggcaaagg 1320
aagcctggga gaatcacagg ttgaggaatg attctgtgat tgtggatact ttccatggcc 1380
tcttcaaatc tactttggtt tgcccagaat gtgctaaggt ttctgtgacc tttgacccat 1440
tttgctatct aacgctgcca ctgcccttga agaaagatcg agttatggag gttttcctgg 1500
ttcctgctga ccctcactgc agacctactc agtaccgtgt gactgtgccg ctgatggggg 1560
ctgtgtccga cctgtgcgag gctctctcca ggctgtctgg cattgctgca gaaaatatgg 1620
tggtcgcaga tgtgtataat caccgattcc acaaaatttt ccaaatggat gaaggtttaa 1680
accacatcat gcctcgggat gacattttcg tgtacgaggt ctgcagcact tccgtggatg 1740
gctcggaatg tgtcacgctt ccagtctact tcagggagag gaagtccagg ccatcaagca 1800
cttcctccgc atcagcgcta tatgggcagc cactattgct ttctgtcccc aagcacaagt 1860
taacccttga gtctttgtac caggctgttt gtgatcgtat cagccgctat gtgaaacagc 1920
ctttacctga tgagtttggc agctcaccct tggagccagg ggcctgcaat ggctccagga 1980
acagctgtga aggagaagat gaggaagaaa tggagcatca ggaagaaggc aaagagcagc 2040
tttcagaaac agaaggcagt ggggaagatg agccaggaaa tgaccccagt gagaccaccc 2100
aaaagaagat caaaggccag ccctgcccaa aaaggctttt taccttcagt cttgtgaact 2160
cctatggaac agctgacata aattcacttg cagctgatgg aaaactactt aaactcaact 2220
ctcgatctac actggccatg gattgggaca gtgaaactcg gagactttac tatgatgagc 2280
aagaatctga ggcctacgag aagcatgtga gcatgttgca gcctcagaag aagaagaaga 2340
ccacagtggc cctgagagac tgcatcgagc tcttcaccac catggagacc cttggggagc 2400
atgaccctg gtactgtccc aactgtaaga agcatcaaca ggccacaaaa aagtttgacc 2460
tatggtcctt gcccaagatc ctggtggtcc acctcaaacg tttctcctac aacagatact 2520

ggagggataa gctcgacaca gtcgtagaat tcccaatcag agggctgaac atgtccgagt 2580
ttgtctgtaa cctgtcagca aggccttatg tgtacgacct cattgccgtg tccaatcatt 2640
atggagccat gggggttggc cactactaaa gcagcttatg tgctatttta ccaacgtcga 2700
gatgatgaat tttataagac accttcactt agcagttctg gttcctctga tggagggaca 2760
cgaccaagca gctctcagca gggctttggg gatgatgagg cttgcagcat ggacaccaac 2820
taatgctgac tccacgatcc tgccaccctg tagcgccagt gtaatccccc aggagaacat 2880
ctttgacact ctgcagactg ctagtgttct gtctaaaaac cagacaagga aatacccttc 2940
ttttatgagc agaaggaaac aaaaaaaaaa aagaagaccg tttacctaga agaagctatg 3000
tcaagaggct gaattatttt tatttttaaa caggtggtga gaaatttctg tgaaacctgt 3060
gaagctgaaa agggggtggg atgggggtac tcaatggagt atgtctgatg gatcccgaag 3120
aatggagagg aacacaggcg ctgagtatgg agcagcctgg agacccaccc cacctgcagg 3180
gctgccctga gcgcctggat ttctggttct gatgccaata ctagtcaccc caagtagtct 3240
gctcacagta acccaaacct caagtaaact ccccttccc tcgctatgtg cactgatatg 3300
ggtttataat tttctgaaag ttacccactg aagcccattt ctccgttggg tcatcttgta 3360

<210> 93
<211> 13855
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507771CB1

<400> 93
ctccgcccct cactgcttcg ttaggttgtg cgtcatttgc ttttcgcttc gcgtagggtg 60
aagctgtagc tacttcggct ttggtgggag ggaggagggg tctggaaagg gctgggctca 120
ggcttttccc gtccggtaga ggttctcgcg ggatcgcgcg gaggcggcgg tggctcggtt 180
actgactgca gtagccgcgg aggcggcggc aaggtcttcg gctcgtgccg ttctaagccg 240
gagaattctc gcgggagcag ggttacgtcc tcgtgggatt cgttggcggt ggctgaggtc 300
ctccagcagc ctgacctgag tgggttagtg atccagagaa accagcaggc caacttggtc 360
aggaaggttc gggaagctgt tggagcagtg tggggaattt cccaccagga tgagtatgat 420
tggctgtgat tttagatcgt aaagctgaaa attgaaatca tgaaagtaga caggactaaa 480
ctgaagaaga cacctactga ggctcctgca gactgcagag ccttaataga caaactcaaa 540
gtttgtaatg atgagcaact tctcttggaa ctgcagcaga tcaaaacatg gaacattgga 600
aagtgcgagt tatatcactg ggtggacctg ttggaccgct tcgatggaat actggcagat 660
gctggacaga cagtggagaa tatgtcatgg atgctcgtat gtgataggcc agaaagagag 720
caactgaaaa tgcttctctt ggctgtgttg aacttcacag ccttgctcat tgagtacagc 780
ttttcccggc atctgtacag ttccatagag catttgacaa ctttattggc ttcctctgat 840
atgcaagtgg tgctggcagt cctcaatctc ctatatgtat ttagcaaaag atcaaactac 900
atcactcgtc tgggatctga caagaggacc ccgctgctaa ctcggctaca acatttggca 960
gagagctggg gtggaaagga gaatggcttt ggacttgcag aatgttgcag agacttgcat 1020
atgatgaaat atccacccag tgcaactaca ctacactttg aattctatgc agatcctggg 1080
gccgaggtca aaattgagaa aaggacaact agtaacacac tacattatat tcacatagag 1140
caacttgaca agatttcaga aagcccttct gaaatcatgg aatctcttac caaaatgtac 1200
agcactccta aggataagca gatgctgtta tttacacaca tacgactggc ccatggcttt 1260
tctaatcaca ggaagcgatt gcaggcagtt caggccagac tgcatgcaat atctatatta 1320
gtgtattcca atgccttgca ggaatcagca aacagtatct tgtataatgg cttgatagag 1380
gagttggtag atgtccttca gataacggat aagcagctta tggagattaa agcagcttct 1440
ttacgaacat taacatcaat tgtccacttg gagagaactc ccaaactcag cagtattatt 1500
gactgtactg gaactgcctc ctaccatgga tttttgccag tgcttgtaag gaactgtatc 1560
caggccatga ttgatccttc catggatcca taccctcacc agtttgccac tgctctcttc 1620
tctttttat accatctggc cagctacgat gctggtggtg aagccttggt ctcctgtgga 1680
atgatggaag ccttattgaa ggtcataaag tttcttggcg atgaacagga ccagataaca 1740
tttgtcacca gagccgtcag agtggttgac cttatcacca acctggatat ggcagctttt 1800
caatcccata gtggactttc tatcttcatt tatagacttg agcatgaagt agatttgtgc 1860
cgaaaagaat gtccgtttgt gatcaagcca aagatccaga gacccaatac tacacaagaa 1920
ggagaggaaa tggaaactga tatggatgga gtccagtgta ttccacaacg agcagcactt 1980
ctgaaatcca tgttgaattt cctcaagaag gccatccaag accctgcttt ctcagatggc 2040
atacgacatg tgatggatgg ttctctgcct acctccctga aacacatcat cagcaatgca 2100
gaatactatg gcccatcact cttcctccta gctactgaag tggtgactgt gtttgtattt 2160
caagaaccat cactgctctc ctcactccag gacaatggat tgacagatgt catgctgcat 2220
gcactgctta tcaaagatgt tcctgctacc cgtgaagtcc ttggctccct cccaaatgta 2280

ttcagtgcac tctgtttgaa tgcccgaggt cttcagtctt ttgttcagtg tcagcctttt 2340
gaacgcctct tcaaagttct tctgtctcca gattacctcc cagccatgcg gaggaggaga 2400
agttctgatc cccttgggga tactgcatcc aacctgggga gtgctgtcga tgagctcatg 2460
agacatcagc ccacccttaa aacagatgca acgactgcca tcatcaagtt acttgaagaa 2520
atctgtaatc ttggaaggga ccccaaatac atctgtcaga agccatcaat ccagaaggca 2580
gatggcactg ccactgctcc tcccccaagg tctaatcatg ccgcagaaga agcctctagt 2640
gaggatgagg aggaagagga agtacaggcc atgcagagct ttaattctac ccagcaaaat 2700
gaaactgagc ctaatcagca ggttgttggt acagaggaac gtattcctat tcccctcatg 2760
gattacatcc ttaatgtgat gaaatttgtg gaatctattc tgagcaacaa tacaacagat 2820
gaccactgcc aggaatttgt gaatcagaaa ggactgttgc ctttggttac cattttgggt 2880
cttcccaatc tgcccattga ctttcccaca tctgctgcct gtcaggctgt tgcaggtgtc 2940
tgcaaatcca tattgacact gtcacatgaa cccaaagtcc ttcaagaggg tctccttcag 3000
ttggactcca tcctctcctc cctggagccc ttacaccgcc ccattgaatc ccctgggggc 3060
tcagtgttgt tgcgagaact ggcttgcgca ggcaatgttg ctgatgctac cctctcagcc 3120
caggccacac ctctgctgca tgcactcact gctgcccatg cctacatcat gatgtttgtt 3180
catacttgca gagttggaca gagtgaaatt cgttccatct ccgtaaacca gtggggctct 3240
caattgggtc tgagtgtttt gagcaagctg agccagttat actgttccct ggtgtgggaa 3300
agcactgtcc tcctctctct gtgtacccca aacagcctac catctgggtg tgaatttggc 3360
caggcagata tgcagaaact ggttccaaag gatgagaagg caggtacgac ccagggcgga 3420
aaaagatcag atggggaaca ggatggagca gctggaagta tggatgcttc tacccagggc 3480
ttattagaag gcattgggct agatggtgac acattggctc ccatggagac agatgaacct 3540
actgcttcag actctaaggg caaatctaaa atcacaccag caatggctgc cagaattaag 3600
caaatcaagc ctttgttatc agcttcctcc agattaggcc gagcacttgc tgagctattt 3660
ggacttcttg ttaaactttg tgtgggatct cctgtccgcc agagaaggag ccatcatgct 3720
gccagcacca ctacagcacc gacacctgcc gcgcgatcaa cagcctcagc tctcactaag 3780
ctcttgacta aggggttatc ttggcagccc ccaccatata cacctactcc ccgattcagg 3840
ctgacattct tcatctgttc agttggtttc acatccccaa tgctgtttga tgagaggaag 3900
tatccctacc acctcatgct gcaaaaattt ctctgctccg gaggccacaa tgctcttttt 3960
gaaactttca actgggctct gtccatggga ggtaaagttc ctgtttctga gggattggaa 4020
cactcagact tgcctgatgg cacaggagaa ttcctagatg cctggcttat gctggtggag 4080
aagatggtga atcccaccac ggtgcttgaa tctccacatt cgctgcctgc caaattgcct 4140
ggaggtgtcc agaactttcc ccagttcagt gcactgcgct tccttgtggt aactcagaaa 4200
gcagccttta cttgcatcaa aaacttatgg aaccggaaac ccctgaaggt atatggtgga 4260
cgaatggctg aatcgatgct ggccattcta tgccacatcc tccgaggaga acctgtgatt 4320
cgagagagac taagcaagga gaaggagggg tctcgaggag aagaggatac agggcaagag 4380
gaaggtggct cccgccggga acctcaagtc aaccagcaac aactgcaaca gctcatggac 4440
atgggcttca caagggaaca tgcaatggag gcactgttga acaccagcac catggagcag 4500
gccacagagt accttttaac ccaccctcct ccaatcatgg gaggagttgt tcgggatctc 4560
agcatgtctg aagaggacca gatgatgaga gcaattgcta tgtctctggg acaggatatt 4620
ccaatggatc aaagggcaga gtcacctgag gaagttgctt gccggaagga ggaagaggaa 4680
cggaaagctc gggaaaagca ggaggaggaa gaggctaaat gtctagagaa gttccaggat 4740
gctgacccgt tggaacaaga tgagctccac actttcacag atactatgtt gccaggctgc 4800
ttccaccttc ttgatgagct gccagacaca gtataccgtg tgtgtgacct gatcatgaca 4860
gcaatcaaac gtaatggagc agattatcgt gacatgattc tgaagcaagt agtcaatcag 4920
gtgtgggaag ctgctgatgt attgatcaaa gctgctcttc ccctgacaac aagtgacaca 4980
aaaaccgtgt cagagtggat aagtcagatg gccacactgc cccaggcctc caatttggct 5040
actagaatct tgcttttaac gctacttttt gaggagttga agctaccttg tgcttgggtg 5100
gttgaatcaa gtggcatcct taatgtccta atcaaactct tggaagtggt tcagccctgc 5160
ctccaggcag ccaaggagca gaaggaagtc cagaccccaa agtggatcac accagtgttg 5220
ctcctgattg atttctatga aaagacagcc atctcctcaa aaaggagagc ccagatgact 5280
aagtacctgc aatccaacag caacaactgg cgctggtttg atgatcgctc tgggcgttgg 5340
tgtagttaca gtgcaagcaa caatagcact attgattctg cctggaaatc tggagagaca 5400
agcgtgcgat tcactgcagg ccgaagaaga tacacggtcc aattcactac aatggtgcag 5460
gttaatgagg aaacagggaa ccgacgccct gtgatgctga ctctcctcag ggtacctcgg 5520
ctgaataaaa attcaaaaaa cagcaatgga caggaactag agaagacgct ggaagaaagc 5580
aaagaaatgg atatcaaacg taaagaaaat aaaggcaatg ataccccttt ggccctagag 5640
agtacaaaca ctgaaaagga gacaagcctg gaggaaacaa aaatcgggga gatcctgatc 5700
cagggcttga cagaagatat ggtgactgtt ttaatccggg cctgcgtgag catgctggga 5760
gtccctgtgg acccagatac tttgcatgcc acccttcgtc tctgtctgag gctcacccgg 5820
gaccacaaat atgccatgat gtttgcagaa ctgaagagta cccgcatgat cttgaatttg 5880
acccagagct caggcttcaa tgggtttact cccctggtca cccttctctt aagacacatc 5940
attgaggacc cctgtaccct tcgtcatacc atggaaaagg ttgttcgctc agcagctaca 6000
agtggagctg gtagcactac ctctggtgtt gtgtctggca gcctcggctc tcgggagatc 6060

aactacatcc ttcgtgtcct tgggccagcc gcatgccgca atccagacat attcacagaa 6120
gtggccaact gctgtatccg catcgccctt cctgcccctc gaggctcagg aactgcttca 6180
gatgatgaat ttgagaatct tagaattaaa ggccctaatg ctgtacagct ggtgaagacc 6240
accccctttga agccctcacc tctgcctgtc atccctgata ctatcaagga agtgatctat 6300
gatatgctga atgctctggc tgcataccat gctccagagg aagcagataa atctgatcct 6360
aaacctgggg ttatgaccca agaggttggc cagctcctgc aagacatggg tgatgatgta 6420
taccagcagt accggtcact tacgcgtcag agcagtgact ttgatacgca gtcaggtttt 6480
tccattaata gtcaggtctt tgctgcagat ggtgcctcca ctgagacttc cgcatctggg 6540
acctcccaag gagaggcttc aactccagag gagtctcgag atgggaagaa agataaagaa 6600
ggggaccggg cctctgagga aggcaaacag aaaggcaagg gcagcaaacc tttaatgcct 6660
acctccacta tccttcgtct tctggcagag ttggtgaggt cctatgttgg tattgctacc 6720
ctgattgcca actacagcta cactgtgggc cagtctgaac tgatcaaaga ggactgcagt 6780
gtgctagctt ttgttctgga ccacctgctc ccacataccc agaatgcaga agacaaggac 6840
acccctgcct tggcccgcct gttcctcgca agcctggctg ctgcagggag tggcacagat 6900
gcccaggtgg ccctagtgaa tgaagtaaaa gcagcccttg gacgggcact ggctatggct 6960
gagagtacag agaaacatgc caggcttcag gcagtgatgt gtatcatcag tactatcatg 7020
gagtcctgcc cctccacctc cagcttctac agcagtgcca cagcgaagac ccagcacaat 7080
ggcatgaaca acatcattcg gcttttcctg aagaagggac tggttaatga cctggccaga 7140
gtacctcaca gcttagacct gtccagtccc aacatggcca acacagtcaa tgctgctctg 7200
aagcctttgg aaacactttc ccggattgtg aaccagccca gtagcctttt tggcagcaag 7260
agtgcttcta gcaagaacaa gtctgagcag gatgcccaag gagcctctca agattccagt 7320
agcaaccagc aggacccagg cgagcctggg gaagcagaag tgcaggagga ggatcatgat 7380
gtcactcaga cagaggtggc agatggggat atcatggatg gggaggctga aaccgactca 7440
gtggtgattg ctgggcagcc tgaggtgctc agttcacaag agatgcaggt tgagaatgag 7500
ctggaggacc tgatagatga gttgcttgag agggatggcg gatctgggaa cagtacaatt 7560
atagtgagca gaagtggaga ggatgaatca caagaggacg tgctgatgga tgaagctcct 7620
tccaacctca gccaagcttc caccttgcag gccaaccgag aagattccat gaatatcctg 7680
gaccctgagg atgaggagga gcacactcag gaagaggaca gcagtggcag taacgaggat 7740
gaggatgata gtcaggatga agaggaggag gaggaggaag atgaggaaga tgatcaggag 7800
gatgatgaag gtgaagaggg agatgaagac gatgacgacg atggctctga gatggaattg 7860
gatgaggatt atcctgatat gaacgcttct cccttggtcc gatttgagcg ctttgaccgg 7920
gaggatgatc tcatcattga gtttgacaac atgttctcca gtgctacaga catccccccca 7980
tccccaggaa atatccctac cacccatcca ctgatggtgc gccatgcaga ccacagttct 8040
ctgacactgg gcagtggctc ttcaacaact cgtctcaccc agggcatcgg gcgcagtcag 8100
aggaccctaa ggcagctgac ggccaatact ggccacacca ttcatgttca ctaccctggg 8160
aatcgccagc ccaaccctcc tcttatactg cagaggttgc ttggtccctc agctgctgct 8220
gacatccttc agctgagcag cagccttccc ctacaaagcc ggggtcgggc ccgcctcctg 8280
gtaggcaacg atgacgtcca catcatcgcc cgttctgatg atgagctgct ggatgacttt 8340
ttccatgatc agagcacagc taccagccaa gcaggaaccc tgtccagcat ccccacagcc 8400
ctgacccgct ggacagaaga atgcaaagtt ctcgatgctg agagcatgca tgactgtgtt 8460
tcagtggtta aagtgtccat tgtcaatcac ctggaattcc tgagggatga ggagctggaa 8520
gaaaggcgag agaagcgcag gaaacaactg gctgaggaag aaacaaagat aactgataaa 8580
ggcaaagaag ataaggagaa cagggatcag agtgcacagt gtactgcatc taagtcaaat 8640
gactccactg aacagaatct ctcagatggg acgcctatgc ctgacagcta cccaacaacc 8700
ccatcttcaa ctgatgcagc tacatctgag tccaaggaga cccttggcac tctgcaatcc 8760
tcacaacagc aaccaacact cccaacccca ccagctttgg gagaggttcc tcaggagctg 8820
cagtctccag ctggagaagg gggcagctct acacagctat tgatgcctgt agagccagag 8880
gaattgggtc ccacaaggcc aagtggggaa gcagaaacaa ctcagatgga gttatcccca 8940
gctcccacta taacctcact ttccccagag agagctgagg attctgatgc actgacggct 9000
gtcagcagtc agctagaagg ctctcctatg gatacaagca gcctggcttc ctgtacctta 9060
gaggaggctg tgggtgacac ttcagcagct ggcagttctg agcagcccag agcaggcagc 9120
tccactcctg gggatgcccc accagctgtg gcggaagtgc aaggcaggag tgatgggtca 9180
ggggaatctg cccagccacc tgaggacagc tccccacctg catcctctga gagctcttcc 9240
accagagatt ctgccgtggc catttctgga gcagattccc gaggaatcct agaagagccg 9300
ttgccttcaa caagcagtga agaagaagat ccccttgcgg gtatcagtct ccctgaaggt 9360
gtggacccct cttttctggc tgccctgcct gatgacatcc gtcgggaagt tctacagaac 9420
cagctaggca ttcgtccacc aacccggact gccccctcca caaatagctc agcgcctgca 9480
gtggtgggga atcctggtgt gactgaagtg agccctgagt ttctggctgc cctgcctcca 9540
gccattcagg aggaagtact ggcacagcag agagctgagc agcagcgacg agaactagca 9600
cagaatgcca gctcagacac ccctatggac cctgtgacct tcatccagac tctgccctca 9660
gacctgcgcc gtagtgtcct agaggatatg gaggacagtg tgttagctgt gatgccacct 9720
gacattgcag ctgaggctca agccctgaga cgagagcaag aagcccggca gcgacagctc 9780
atgcatgagc gtctgtttgg gcacagtagc acctccgcac tctctgctat tctccgaagc 9840

ccggctttca ccagtcgctt aagtggcaac cgtggggtcc agtatactcg ccttgctgtg 9900
cagagaggtg gcaccttcca gatgggggt agcagcagcc ataacaggcc ttctggcagt 9960
aatgtagata ctctcctccg cctccgagga cggctccttc tggaccacga agccctttct 10020
tgtctcttgg tcctactttt tgtggatgag ccaaagctca atactagccg tctacaccga 10080
gtactgagaa atctctgcta ccatgcccag acccgccact gggtcatccg cagtctgctc 10140
tccatcttgc agcgcagcag tgagagtgag ctatgcattg aaacacccaa actcactaca 10200
agtgaggaaa agggcaaaaa gtcgagcaag agctgtgggt caagtagcca tgagaaccgt 10260
cccctggacc tgctacacaa gatggagtca aagagctcca accagctttc ctggctctca 10320
gtatccatgg atgcagccct aggctgcagg actaatatat ttcagatcca gcgttcaggg 10380
gggcgtaaac ataccgagaa gcatgcaagc ggtggctcca ccgtccacat ccatccccaa 10440
gctgctcctg ttgtctgcag acacgttttg gatacactca ttcaattggc caaggtattt 10500
cccagccact tcacacagca gcggaccaaa gaaacaaact gtgagagtga tcgggaaagg 10560
ggcaataagg cctgtagccc atgctcctca cagtcctcca gcagtggcat ttgcacagac 10620
ttctgggact tattggtaaa actggacaac atgaatgtca gccggaaagg caagaactcc 10680
gtgaagtcag tgccagtgag cgctggcggt gagggggaaa cctctccata cagcctcgag 10740
gcctctccac tggggcagct catgaacatg ttgtcacacc cagtcatccg ccggagctct 10800
ctcttaactg agaaactcct cagactcctt tctctcatct caattgctct cccagaaaac 10860
aaggtgtcag aagcacaggc taattctggc agcggtgctt cctccaccac cactgccacc 10920
tcaaccacat ctaccaccac caccactgcc gcctccacca cgcccacacc ccctactgca 10980
cccacccctg tcacttctgc tccagccctg gttgctgcca cggctatttc caccattgtc 11040
gtagctgctt cgaccacagt gactaccccc acgactgcta ccactactgt ttcaatttct 11100
cccactacta agggcagcaa atctccagcg aaggtgagtg atgggggcag cagcagtaca 11160
gactttaaga tggtgtcctc tggcctcact gaaaaccagc tacagctctc tgtagaggtg 11220
ttgacatccc actcttgttc tgaggaaggc ttagaggatg cagccaacgt actactgcag 11280
ctctcccggg gggactctgg gacccgggac actgttctca agctgctact gaatggagcc 11340
cgccatctgg gttataccct ttgtaaacaa ataggtaccc tgctggccga gctgcgggaa 11400
tacaacctcg agcagcagcg gcgagcccaa tgtgaaaccc tctctcctga tggcctgcct 11460
gaggagcagc cacagaccac caagctgaag ggcaaaatgc agagcaggtt tgacatggct 11520
gagaatgtgg taattgtggc atctcagaag cgacctttgg gtggccggga gctccagctg 11580
ccttctatgt ccatgttgac atccaagaca tctacccaga agttcttctt gagggtacta 11640
caggtcatca tccagctccg ggacgacacg cgccgggcta acaagaaagc caagcagaca 11700
ggcaggctag gttcctccgg tttaggctca gctagcagca tccaggcagc tgttcggcag 11760
ctggaggctg aggctgatgc cattatacaa atggtacgtg agggtcaaag ggcgcggaga 11820
cagcaacaag cagcaacgtc ggagtctagc cagtcagagg cgtctgtccg gagggaggaa 11880
tcacccatgg atgtggacca gccatctccc agtgctcaag atactcaatc cattgcctcc 11940
gatggaaccc cacaggggga gaaggaaaag gaagaaagac cacctgagtt acccctgctc 12000
agcgagcagc tgagtttgga cgagctgtgg gacatgcttg ggagtgtct aaaggaacta 12060
gaggaatccc atgaccagca tgcggtgcta gtgctacagc ctgctgtcga ggccttcttt 12120
ctggtccatg ccacagagcg ggagagcaag cctcctgtcc gagacacccg tgagagccag 12180
ctggcacaca tcaaggacga gcctcctcca ctctcccctg ccccttaac cccagccacg 12240
ccttcctccc ttgacccatt cttctcccgg gagccctcat ctatgcacat ctcctcaagc 12300
ctgcccctg acacacagaa gttccttcgc tttgcagaga ctcaccgcac tgtgttaaac 12360
cagatcctac ggcagtccac gacccacctt gctgatgggc cttttgctgt cctggtagac 12420
tacattcgtg tcctcgactt tgatgtcaag cgcaaatatt tccgccaaga gctggagcgt 12480
ttagatgagg ggctccggaa agaagacatg gctgtgcatg tccgtcgtga ccatgtgttt 12540
gaagactcct atcgtgagct gcatcgcaaa tcccccgaag aaatgaagaa tcgattgtat 12600
atagtatttg aaggagaaga agggcaggat gctggcgggc tcctgcggga gtggtatatg 12660
atcatctctc gagagatgtt taaccctatg tatgccttgt tccgtacctc acctggtgat 12720
cgagtcacct acaccatcaa tccatcttcc cactgcaacc ccaaccacct cagctacttc 12780
aagtttgtcg gacgcattgt ggccaaagct gtatatgaca accgtcttct ggagtgctac 12840
tttactcgat cctttacaa acacatcttg ggcaagtcag tcagatatac agatatggag 12900
agtgaagatt accacttcta ccaaggtctg gtttatctgc tggaaaatga tgtctccaca 12960
ctaggctatg acctcacctt cagcactgag gtccaagagt ttggagtttg tgaagttcgt 13020
gacctcaaac ccaatggggc caacatcttg gtaacagagg agaataagaa ggagtatgta 13080
cacctggtat gccagatgag aatgacagga gccatccgca agcagttggc ggctttctta 13140
gaaggcttct atgagatcat tccaaagcgc ctcatttcca tcttcactga gcaggagtta 13200
gagctgctta tatcaggact gcccaccatt gacatcgatg atctgaaatc caacactgaa 13260
taccacaagt accagtccaa ctctattcag atccagtggt tctggagagc attgcgttct 13320
ttcgatcaag ctgaccgtgc caagttcctc cagtttgtca cgggtacttc caaggtaccc 13380
ctgcaaggct ttgctgccct cgaaggcatg aatggcattc agaagtttca gatccatcga 13440
gatgacaggt ccacagatcg cctgccttca gctcacacat gttttaatca gctggatctg 13500
cctgcctatg agagctttga gaagctccgc cacatgctac tgttggctat ccaggagtgc 13560
tctgaaggct ttgggctggc ctaataaggc cctgcccaac tccgtggggt tttttttacc 13620

attgttggac ctggggaggg gggagttaaa aaaagaacca gaaagaaatt gtcaaaaacc 13680
aataaatgaa atccaccaac tcaccgtgtg tgtcccagct gccccatctt ccccagcgca 13740
tacctgttcc tcttctcatt ctctccccgc cgcctgtttc ctcaccttct ctcccctttc 13800
catgccgtcc atgatcccca ccccatgtgt tttaaaaagg cagtagcctt tgcag 13855

<210> 94
<211> 1554
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504732CB1

<400> 94
cggagactgg ccgggtagcc ccgcccccgt agttagcgcg gtgcttctct tccgctccgg 60
gtcggctccg tttccctttc cgggcgggca ggcggcggac cccagtgtct ttatccctct 120
tttgcacagt cagcttctgc agctctcccg ggctagcatg gcagcgtgga agagttggac 180
ggccctgcgg ctctgcgcca cagttgttgt acttgatatg gtcgtctgta aaggatttgt 240
agaagattta gatgaatcgt ttaaagaaaa tcgaaatgat gacatttggc ttgtagattt 300
ttatgcgcca tggtgtggcc attgtaaaaa gctggaacca atttggaatg aagttggtct 360
tgagatgaaa agcattggtt ctccagttaa ggttggaaag atggatgcta cttcctattc 420
taattaaaag gggacttggc atataattat agaggaccac gaacaaaaga tgatattatt 480
gagtttgctc acagagtatc tggggctcta attcggccac ttccaagtca acaaatgttt 540
gaacatatgc agaagagaca ccgtgtattt ttcgtttatg taggtggaga atcacctttg 600
aaagagaaat acatagatgc tgcttcagaa ttgattgtat atacatactt cttttctgcc 660
tcagaagaag tggttcctga gtatgtgaca ctaaaagaga tgccagctgt gcttgttttc 720
aaagatgaaa cttactttgt ttatgatgag tatgaagatg gtgatctgtc atcatggatc 780
aacagggaaa ggtttcagaa ttaccttgct atggatggct tcctcttgta tgaacttgga 840
gacacaggaa agcttgtggc tcttgcagtt attgatgaga aaaatacatc agttgaacat 900
accagattga agtcaattat tcaggaagtt gcaagagatt acagagacct cttccatagg 960
gattttcagt ttggccacat ggatggaaat gactacataa ataccttgct gatggatgaa 1020
ttgacagtcc caactgtagt tgtactgaat acttcaaacc agcaatattt cttgctagat 1080
agacagatta agaatgttga agacatggtc cagtttatta ataacatttt ggatggcaca 1140
gtagaagccc aaggaggtga tagcattttg cagagattga aaagaatagt atttgatgcc 1200
aaatctacta ttgtgtctat attcaagagc tcaccactga tgggctgctt tctctttggc 1260
ctgccactgg gtgtcatcag tatcatgtgc tatggaatct acacagccga cacagatgga 1320
ggttatatag aagaacgata tgaagtgtct aaaagtgaaa atgaaaacca agaacagata 1380
gaagagagca aagaacagca ggagcccagc agtggaggat ctgtagtgcc tacagtgcag 1440
gagcccaagg atgtattaga aaagaagaaa gattgagact tgatgactat aaaatatttg 1500
ttaggacttc aaattattaa agagtctatt tattgaattt agacatttaa tctg 1554

<210> 95
<211> 2142
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 950917CB1

<400> 95
cgcattttac actccctttg gtttctaccc cttcgctttt tggtgcccct ttaatctctg 60
gaccccttaa ctcccaacgt tctgtggaaa attttttttt tttttttttt tttttttca 120
atcatcctct ttatagatga aatataataa cagatactag tgagattgtt gggatgatta 180
actgcaaatg tacaaagtcc tttttaacat agtagataat tatttattat tattcctgat 240
agacacagca cagctataag ctgactatat accaagaaca atgagtagga ttacctctta 300
catttctttt cccaacataa tgctttgctc ctgatctgac cagggctttg ctaaagtcac 360
acatccagta agtggccgaa gcagatcttc agacccacag tttggctgcc aaaggaccac 420
agtgagtatt tctaatgctg aaacaagagg gtgcagctct gcagcccaga cggctgggcc 480
tcagggtagg taaccaagac agctcaatgc atctctccct cagccgtccc tagcccctaa 540
aaacagagct gggggtccca gtcctctgtg ggacatggcc ctctttcatc accagcttcc 600
agcctaccct gcatctcttc cctctgaatt gcaggaaccc tcccagcctc catgccgcct 660

ttcctgatca ccctcttcct ctttcactct tgctgcctcc gagcaaatgg ccacctccgt 720
gaaggaatga cattgctgaa gactgagttt gcacttcacc tctaccagag tgtggccgcg 780
tgtagaaatg agacgaactt tgtcatctct cctgctggtg tgtccctccc cctggagatc 840
ctgcagtttg gagcagaagg gagcactggt cagcagctgg cagatgccct ggggtacact 900
gtccatgaca aaagggtgaa agatttcttg catgctgttt atgccacact acccacctcc 960
agccaaggca ccgagatgga gctggcctgc agcctttttg tgcaagtggg aacgccactg 1020
tccccctgct ttgtggagca cgtctcctgg tgggctaaca gcagcctgga accagccgac 1080
ctcagtgagc ccaatagcac cgccatccag actagcgaag ggcctccag agagactgca 1140
ggtgggggcc ccagtgaggg ccctggtggc tggccgtggg agcaagtcag tgcagcattt 1200
gctcagcttg tgcttgtgag caccatgtcc ttccaaggca cttggcgaaa gagattctcc 1260
tccacagaca cacagatcct gcctttcacc tgtgcctatg gcctcgtcct tcaggtcccc 1320
atgatgcacc aaacgaccga ggtcaactac ggtcagttcc aggacactgc aggccatcag 1380
gtgggggtgc tggagcttcc ttacctggga agtgcagtga gtctgttcct ggtgctgccc 1440
cgtgacaaag acacccccct gagccacatc gagccacacc tcacagccag caccatccac 1500
ctctggacca ccagcctgag gagagccagg atggatgtgt tcctgcccag gtttaggatc 1560
caaaatcaat tcaacttaaa aagcatttta aattcttggg gagtcaccga tctttttgat 1620
ccactcaaag ctaacttgaa aggaatttca ggccaagatg gcttttatgt ttctgaagca 1680
atccacaagg ccaagattga agttttggag gaaggcacca aggcatctgg agccacagct 1740
ctgttgttat tgaaaaggtc tcggattcct atttttaaag cagatcggcc attcatctat 1800
ttcctgagag aacctaacac aggtattaca gtattttttg acaggattca gataatttat 1860
cagtgtctct ctagcaacaa gggctcattt gtccactatc ccctcaaaaa taagcattct 1920
ttctagctgt ttttatggtg cttcagaagc tacccagttg ctatcttcta tcctaagaaa 1980
ggggactatt tcgtaagtac aaaggaaact tatcctccat gtttttacc ttctcattct 2040
cctcagctca tatatagtta atgtaaattt ttttttttaa tgctataggt ttctacttct 2100
gaagacactt ggaaaaatca tgaaagtaaa ataaattata aa 2142

<210> 96
<211> 2637
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7459720CB1

<400> 96
ggactataga agagcccagt gtgctggaaa gggggccggt ctgtagcctt tgctgagtca 60
acactactca cagtgaaagt gatgcctttt ctccatttgc acaacatgag tgtgacgtgg 120
ttagccagac caggagctca acctcatgta gaatccagtc cactgctctg atgccgaagg 180
ggaggcagaa agtgccacac ttggatgccc ccctgggcct gcccacctgc ctctggctgg 240
aattagccgg gctcttctta ctggttccct gggtcatggg cctggcaggg acaggtgggc 300
ctgatggcca gggcacaggg gggccgagct gggctgtgca cctggaaagc ctggaaggtg 360
acggggagga agagactctg gagcagcagg cggatgcctt ggcccaggca gcagggctgg 420
tgaatgctgg acgcatcgga gagcttcagg ggcactacct ctttgtccag cctgctgggc 480
acaggccggc cctggaggtg gaggccatcc ggcagcaggt ggaggctgtg ttggctgggc 540
atgaagctgt gcgctggcac tcagagcaga ggctgctaag gcgggccaag cgcagcgtcc 600
acttcaacga ccccaagtac ccgcagcaat ggcacctgaa taaccgacgg agcccgggca 660
gggacatcaa cgtgacgggt gtgtgggaac gcaatgtgac tgggcgaggg gtgacggtgg 720
tggtagtgga tgacggagtg gaacacacca tccaggacat tgcacccaac tatagccctg 780
agggtagcta tgacctcaac tctaatgacc ctgaccccat gccccacccg gatgtggaga 840
atggcaacca ccatggcacg cgatgtgcag gagagatcgc ggctgtgccc aacaacagct 900
tctgtgccgt gggcgtggcc tacgggagcc gcatcgcagg tatccgggta ctggatggac 960
ctctcacaga cagcatggag gcagtggcgt tcaacaagca ctatcagatc aatgacatct 1020
acagctgcag gtatttattt ctcaacagct ggggaccaga tgacgatggg aagacagtgg 1080
atggccccca tcagcttgga aaggctgcct tacaacatgg ggtgattgct ggtcgccagg 1140
gctttgggag catctttgtg gtagccagtg gcaacggagg ccaacacaac gacaactgca 1200
actacgatgg ctacgccaac tccatctaca ccgtcaccat aggagctgtg gatgaggagg 1260
gacgcatgcc tttctatgca gaagaatgtg cctccatgct ggcagtcacc ttcagtggtg 1320
gggacaagat gcttcggagc attgtgacca ctgactggga ccttcagaag ggcactggct 1380
gcactgaggg ccacacaggg acctcagctg cagcgcctct ggcagctggc atgatagcct 1440
taatgctgca ggtgcggccc tgcctcacgt ggcgtgacgt ccagcacatc attgtcttca 1500
cagccacccg gtatgaggat cgccgtgcag agtgggtcac caacgaggca ggcttcagcc 1560
atagccacca gcacggtttc ggcctcctca acgcctggag gctcgtgaat gcagccaaga 1620

tctggacatc tgtcccttac ttagcatcct acgtcagtcc cgtgttaaaa gaaaacaagg 1680
cgattccgca gtccccccgt tccctggagg tcctgtggaa tgtcagcagg atggacctgg 1740
agatgtcagg gctgaagacc ctggagcatg tggcagtgac agtctccatc actcacccac 1800
ggcgcggcag cttggagctg aagctgttct gccccagtgg catgatgtcc ctcatcggcg 1860
ccccccgcag catggactcg gatcccaacg gcttcaatga ctggaccttc tccactgtgc 1920
gatgctgggg ggagagagcc cgagggacct acaggcttgt catcagggat gtcggggatg 1980
agtcattcca ggtcggcatc ctccggcaat ggcagctgac cctatatggc tctgtgtgga 2040
gtgcagtaga catcagggac agacaaaggc tgttagagag tgccatgagt ggaaaatacc 2100
tgcacgatga cttcgccctg ccctgcccac cggggctgaa aattcctgag gaagatggtt 2160
acaccatcac ccccaacacc ctcaagaccc tggtgctggt aggctgtttc accgtcttct 2220
ggactgttta ctacatgctg gaagtatatt tgagccagag gaatgtggct tccaatcaag 2280
tttgtaggag tggaccctgc cactggcccc atcggagccg gaaagccaag gaggaaggga 2340
cagagctaga atcagtgcca ctttgcagca gcaaggatcc agacgaagtg gaaacagaga 2400
gcaggggccc tcccaccacc tctgacctcc ttgccccaga cctgctggag caaggggact 2460
ggagcctgtc ccagaacaag agcgccctgg actgccctca tcagcaccta gacgtaccgc 2520
acgggaagga ggagcagatc tgctgacctc agggcctgac agtgtgggac aggctcttct 2580
ttcccaaatt agggagctct tgacagaaag cagttctgat gcttacatct ggaatct 2637

<210> 97
<211> 2447
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503300CB1

<400> 97
agtgcccacg tcctgagccg cgcccagact cggaagcggg ccgtcacgtg gggcggcgtg 60
ggctgctgcc ttggaaacgc ggagcgcgcg ctcccagcga aagcagcagg gcagggatct 120
gcgttggagg aagggactgc tctggtgcta gaatgctgtg cgtcggaagg ctgggcggct 180
tgggagccag agcagcagct ctgccgcccc gccgggcggg ccggggaagc ctcgaagccg 240
ggatccgggc ccgaagggtc agcaccagct ggtctcccgt gggcgccgcc ttcaatgtca 300
agccccaggg cagccgcttg gacctgttcg gcgagcgccg gggtcttttt ggagttcctg 360
agctgagtgc cccagaagga tttcatattg cacaagaaaa agccttgaga aagacagaat 420
tgcttgtgga ccgtgcatgt tccaccccac ctgggcccca gaccgtgctg atcttcgatg 480
agctctcgga ttccttatgc agagtggccg acttggctga ttttgtgaaa atcgctcacc 540
ctgagccagc attcagagaa gctgcggaag aagcttgtag aagtattggc accatggtag 600
agaagttgaa cacaaatgtg gatttatatc aaagtttgca aaaattacta gctgataaaa 660
aacttgtgga ttcccttgat ccagaaacaa ggcgagtggc tgaactgttt atgtttgatt 720
ttgaaattag tggaatccat ctagacaaag aaaagcgtaa aagagcagtg gacctcaatg 780
ttaaaatctt ggatttgagt agtacatttc ttatgggaac caattttccc aacaagattg 840
agaagcatct cttaccagaa cacattcgtc gtaactttac atctgctggg gatcatatca 900
taattgatgg tctccacgca gaatcaccag atgacttggt gcgagaagct gcttataaaa 960
tttttcttta tcccaatgct ggtcaattga aatgtttaga agaattgctc agcagcagag 1020
atcttctggc aaagttggtg gggtattcca cgttttctca cagggctctc caaggaacga 1080
tagctaaaaa tccagagact gtcatgcagt tccttgaaaa actatctgac aaactttctg 1140
aaagaactct gaaagatttt gagatgatac gagggatgaa aatgaaactg aatcctcaaa 1200
attccgtata atattgagcc cagcctatat tgcccgtttt tctctcttgg agcatgcatg 1260
gaaggcctga atattttgct taacagactg ttgaggattt cattatatgc agagcagcct 1320
gcaaaaggag aggtgtggag cgaagatgtc cgaaaactgg ctgttgttca tgaatctgaa 1380
ggattgttgg ggtacattta ctgtgatttt tttcagcgag cagacaaacc acatcaggat 1440
tgccatttca ctatccgtgg aggcagacta aaggaagatg gagactatca actcccagtt 1500
gtagttctta tgctgaatct tccccgttcc tcaaggagtt ctccaacttt gctaactcct 1560
ggcatgatgg aaaatctttt ccatgaaatg ggacatgcca tgcattcaat gctaggacgt 1620
actcgttacc aacacgtcac tgggaccagg tgccctactg attttgctga ggttccttct 1680
attctgatgg agtactttgc aaatgattat cgagtagtta accaatttgc cagacattat 1740
cagactggac agccactgcc aaaaaatatg gtgtctcgtc tttgtgaatc taaaaaggtt 1800
tgtgctgcag ctgatatgca acttcaggtc ttttatgcca ctctggatca aatctaccat 1860
gggaagcatc ccctgaggaa ttcaaccaca gacattctca aggaaacaca agagaaattc 1920
tatggcctac catatgttcc aaatactgcc tggcagctgc gattcagcca cctcgtgggg 1980
tatggtgcta gatattactc ttacctcatg tccagagcgg tcgcctccat ggtttggaag 2040
gagtgttttc tacaggatcc tttcaacagg gctgccgggg agcgctatcg cagggagatg 2100

ctggcccacg gtggaggcag ggagcccatg ctcatggttg aaggtatgct tcagaagtgt 2160
ccttctgttg atgacttcgt aagtgccctc gtttccgact tggatctgga cttcgaaact 2220
ttcctcatgg attctgaata aaagaaacac tctacacctc ttaaatcaag gtcatgtaga 2280
taatgacttt gttataaatg ctacagctgt gagagcttgt ttctgatttc attgttcgct 2340
tctgtaattc tgaaaaactt taaactggta gaacttggaa taaataattt gttttaatta 2400
aaaaaaaaaa aaaaaaaaaa aaaaagggcg ccgctcgcga tctagac 2447

<210> 98
<211> 2384
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503334CB1

<400> 98
tatgaagaag aaactcaaaa taacaggagt ggcttatgga actacatgga ggtaacagag 60
gagggtacca accaaaggcc cttgagcaat caggatgttg gctctggcag gctgccctga 120
gcacgctcaa ccccaacccc acggacagct gtcccctcta cctgaactac gccaccgtgg 180
ctgccctgcc ctgcagggtg agccggcaca acagcccctc ggccgcccac ttcatcacgc 240
ggctggtgcg gacctgcctg ccgcccggag cgcatcgctg cattgtgatg gtctgcgagc 300
agccggaggt ctttgcttcc gcctgtgccc tggcccgggc cttcccgctg ttcacccacc 360
gctcaggtgc ctctcggcgc ttggagaaga agacggtcac cgtggagttt ttcctggtgg 420
gacaagacaa cgggccggtg gaggtgtcca cattgcagtg cttagcgaat gccacagacg 480
gcgtgcggct agcagcccgc atcgtggaca caccctgcaa tgagatgaac accgacacct 540
tcctcgagga gattaacaaa gttggaaagg agctggggat catcccaacc atcatccggg 600
atgaggaact gaagacgaga ggatttggag gaatctatgg ggttggcaaa gccgccctgc 660
atcccccagc cctggccgtc ctcagccaca ccccagatgg agccacgcag accatcgcct 720
gggtgggcaa aggcatcgtc tatgacactg gaggcctcag catcaaaggg aagactacca 780
tgccggggat gaagcgagac tgcggggggtg ctgcggccgt cctgggggcc ttcagagccg 840
caatcaagca gggtttcaaa gacaacctcc acgctgtgtt ctgcttggct gagaactcgg 900
tggggcccaa tgcgacaagg ccagatgaca tccacctgct gtactcaggg aagacggtgg 960
aaatcaacaa cacggatgcc gagggcaggc tggtgctggc agatggcgtg tcctatgctt 1020
gcaaggacct gggggccgac atcatcctgg acatggccac cctgaccggg gctcagggca 1080
ttgccacagg gaagtaccac gccgcggtgc tcaccaacag cgctgagtgg gaggccgcct 1140
gtgtgaaggc gggcaggaag tgtggggacc tggtgcaccc gctggtctac tgccccgagc 1200
tgcacttcag cgagttcacc tcagctgtgg cggacatgaa gaactcattg gcggaccgag 1260
acaacagccc cagctcctgt gctggcctct tcatcgcctc acacatcggc ttcgactggc 1320
ccggagtctg ggtccacctg gacattgctg caccggtgca tgctggtgag cgagccacag 1380
gcttcggtgt ggccctcctg ctggcgctct tcggccgtgc ctctgaggac cctctgctga 1440
acctggtgtc cccactgggc tgtgaggtgg atgtcgagga gggggacctg ggagggact 1500
ccaagagacg caggcttgtg tgagcctcct gcctcggccc tgacaaacgg ggatctttta 1560
cctcactttg cactgattaa ttttaagcaa ttgaaagatt gcccttcata tgggttttgg 1620
tttgtctttc tggtcgtcag cgtggtggtg gaaacagctg aagttttagg agacagctta 1680
gggtttggtg cgggccacgg ggaggggacc gggaagcgct ggggcttgtt tctgtttgtt 1740
acttacagga ctgagacatc ttctgtaaac tgctacccct ggggccttct gcaccccggg 1800
gtgaggcctc ctgcctgcct ggtgccctgt cccagcccca ggtcccgtgc agggcacctg 1860
cgtggctgac agccaggctc ttactccagc cggggctgcc agcgcatcca gccagcccag 1920
ccctgtgaaa gatggagctg acttgctgca ggggacctga tttatagggc aagagaagtc 1980
acactctggc ctctcagaat tcacttgagg ttcaattaaa tacagtcaca ccgccccctc 2040
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaag 2100
agaggggggg ggagcagaaa aggaggaggc gcgcgaaccg gggggaaata taagcggaga 2160
aggagaacag agggggagca caccgagtcc cacaagaagg gggcgcgcaa ataacacacc 2220
ggtcgggcac aaatcaagcg ggcacaaaca cacggcggga ggaaatcgtg tcccacagaa 2280
cgaacgaggg ggcgacgaaa acacacacca gacgctcgcg gagaacccac aaagagggag 2340
cggcgactaa caccagaaaa agacggcgaa acaccggaaa gacg 2384

<210> 99
<211> 778
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503341CB1

<400> 99
gccagtattg ccgggctatg gcggcgagca ctggctacgt gcgactgtgg ggagcggcgc 60
ggtgctgggt gctgcggcgg ccgatgctgg ccgccgccgg ggggcgggtt cccactgcag 120
caggagcgtg gttgctccga ggccagcgga cctgcgacgc ctctcctcct tgggcactgt 180
ggggccgagg cccggcaatt gggggccaat ggcgggggtt ttgggaagcg agcagccgcg 240
gcggaggcgc attctcgggg ggcgaggacg cctccgaggg cggcgcggag gaaggagccg 300
gcggcgcggg gggcagcgcg ggcgccgggg aaggcccggt cataacggcg ctcacgccca 360
tgacgatccc cgatgtgttt ccgcacctgc cgctcatcgc catcacccgc aacccggtgt 420
tcccgcgctt tatcaagatt atcgaggtta aaaataagaa gttggttgag ctgctgagaa 480
ggaaagttcg tctcgcccag ccttatgtcg gcgtctttct aaagagagat gacagcaatg 540
agtcggatgt ggtcgagagc ctggatgaaa tctaccaccc cgggactgca ggcggcggat 600
gtcaggccct gtctgggcca gaactgagcg ctgtggggag cgcgcccgga cctggcagtg 660
gagccaccga gcgagcagct cggtccagtg acccagatcc cagggacctc agtcggctta 720
atcagagtgt ggcatagaag ctatttaatg attaaagtca tttgcagtaa aaaaaaa 778

<210> 100
<211> 3705
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509936CB1

<400> 100
gcctacggat gggaaccacg gactgacctg acctgcgtcc tcggggccac tgcactgggt 60
gattcacgtg tgcccaggcc ctgcaggtgg gccggacctg ggagtggcag ccacacccac 120
tcgcccgggt ctatggccac ggtcggagag tggtcctgtg tgcgctgcac cttcctgaac 180
ccggccggcc agcgccagtg ctccatctgc gaggctcccc ggcacaagcc cgacctcaac 240
cacatcctgc ggctcagcgt ggaggagcag aaatggccct gcgcccgctg caccttccgc 300
aacttcctgg gcaaggaggc ctgcgaggtg tgcggcttca ccccggagcc tgcgcctggg 360
gctgccttcc tgccagtcct caacggggtc ctccccaagc cacccgccat cctgggggag 420
cccaagggca gctgccagga ggaagcaggt ccagtgagga ctgcggggct ggtggccacg 480
gagcccgcca gggggcagtg cgaggacaag gacgaggagg agaaggagga gcaggaggag 540
gaggagggag cggcggagcc cagagggggc tgggcgtgtc cgcgttgcac gctgcacaac 600
acgcccgtgg ccagctcctg ctccgtctgc gggggcccac gcaggctctc gctgccacgg 660
atccctcctg aggccctggt ggtcccggaa gtggtggccc cggccggctt ccacgtcgtg 720
cctgccgcgc ctccacctgg cctccccggg gaaggtgccg aggccaaccc cccagccacc 780
agccagggcc cagctgccga accagagccg cccagggtcc cgcccttcag ccccttctcg 840
tccaccctgc agaacaaccc cgtgccgcgc agccgacgcg aggttccccc ccagctgcag 900
ccaccggtgc ctgaggctgc ccagccgtca ccctctgccg gctgcagggg agccccccag 960
ggctcgggct gggctggggc ctcccgccta gcagagttgc tgtctggcaa gcggctgagt 1020
gtgctggagg aagaggccac ggagggtggc accagccgcg tagaggccgg cagctccacc 1080
tcgggcagtg acatcattga cctggccgga gacaccgtgc gttacacgcc cgccagcccc 1140
tccagccccg acttcaccac ctggtcatgt gccaagtgca cgctcagaaa ccccacagtg 1200
gcccccaggt gctcggcctg cggctgctcc aaactgcacg gcttccagga gcatggcgag 1260
ccccccaccc actgccccga ctgtggggcc gacaagccca gcccctgcgg cagaagctgc 1320
ggacgggtgt cctcggccca gaaggccgcc cgcgtcctgc ccgagcgccc gggccagtgg 1380
gcctgccctg cctgtaccct gctcaacgca ctgcgggcca agcactgcgc cgcctgccac 1440
acgcctcagc tcctggtggc ccagcggcgg ggggccgcgc ccctgaggcg cagggagagc 1500
atgcacgtgg agcagcggcg gcagacagac gagggcgagg ccaaggcact ctgggagaac 1560
atcgtggcct tctgccggga gaacaatgtg agcttcgtgg atgacagctt ccctcccggg 1620
cccgagtctg tcggcttccc cgcgggtgac agcgtgcagc agcgtgtgag gcagtggctg 1680
cgaccccagg agatcaactg ctccgtcttc agggaccaca gggccacgtg gtctgtgttc 1740
cacacactgc ggccctcaga catcctgcag ggctgctgg ggaactgctg gttcctgagc 1800
gccctggcgg tgctggcgga gcggccggac ctggtggagc gggtgatggt cacgcgcagc 1860
ctgtgtgcag agggcgccta ccaggtgcgg ctgtgcaagg acggcacgtg gaccacggtg 1920
ctggtggacg acatgctgcc ctgtgatgag gccggctgcc tcctcttctc acaggcgcag 1980
cggaagcagc tgtgggtggc cctcatcgag aaggcgctgg ccaagctgca cggctcctac 2040

tttgcgctcc aggcgggccg cgccatcgaa ggcctggcca cgctcaccgg cgccccctgt 2100
gagagcctgg cgctgcagct cagctccact aaccccgcg aggagcccgt tgacactgac 2160
ctcatctggg ccaaaatgct gagttctaag gaggctgggt tcctcatggg tgcctcctgt 2220
ggcgggggca acatgaaggt ggacgattcg gcctacgaga gcctgggcct gcgcccccgg 2280
catgcctact ccatcctgga tgtccgagat gtccagggca ccaggcttct gcggctccga 2340
aacccgtggg gccgtttctc ctggaacggc agctggtccg acgagtggcc acactggccg 2400
gggcacctgc gtggcgagct catgccgcac ggcagcagtg agggtgtctt ctggatggag 2460
tacggcgact ttgtcaggta cttcgactcc gtggacatct gtaaggtgca ctcggactgg 2520
caggaggcgc gggtgcaggg ctgctttccc agctcggcca gcgcgcccgt gggggtaaca 2580
gcgctcacgg tgctggagcg ggcctcgctg gagttcgcgc tcttccagga gggcagcagg 2640
cgctcggacg ccgtggacag ccacctgctg gacctgtgca tcctggtgtt ccgggccacg 2700
ttcggcagcg gcggccacct cagcctgggc cgcctcctgg cccacagtaa gcgcgcggtc 2760
aagaagttcg tcagctgcga cgtcatgctg gagcctggcg agtacgctgt ggtgtgctgc 2820
gccttcaacc actgggggcc gcccctgccg ggcacccctg ccccccaggg tacgtggccc 2880
ctaccccagg ctcatgcccc aggcccacgg ggagggctgc ggttcacacg cccgtccttg 2940
tagcctccag cccctcggca ggggtcccga gagcctcccc agagccgccg ggccacgtgc 3000
tggctgtgta cagctcgagg ctggtcatgg tggagcccgt ggaagcccag ccgaccacgc 3060
tggccgacgc catcatcctg ctcaccgaga gccgcggaga gcggcacgag ggccgtgagg 3120
gcatgacctg ctactacctg acacacggtt gggcggggct catcgtggtg gtggagaacc 3180
gacaccccaa ggcctacctg cacgtgcagt gtgactgcac cgacagcttc aacgtggtgt 3240
ccacacgcgg cagcctgcgt acccaggata gcgtgccacc cctgcacagg caggtcctgg 3300
tgatcttgtc ccagctagag ggcaacgccg gcttctctat cacccaccgc ctggcacatc 3360
gcaaggcagc ccaggccttc ctcagtgact ggacagcctc caaggggacc cacagccccc 3420
cgctcacgcc agaggtcgcc ggtctgcatg ggccccgacc gctgtgacca ccatgcctgg 3480
ggcaggggct gtgcacagac ggacccccca cccccacacg cactttatga gggagacccc 3540
gacagaggac gcttgagcca gaatctcagg accccgccca gggagctgcc agcccagggc 3600
tcagcttccc atgggccccc cgccccctc cttcccctcc ctgaacccca cagtccgtct 3660
ggcacaggcc tcctggccgc cacgcagaat acctcgaacc aggtg 3705

<210> 101
<211> 3872
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509986CB1

<400> 101
cccgcgcgtg ccgcggcatg gcggccgtgc gcggggcgcc cctgctcagc tgcctcctgg 60
cgttgctggc cctgtgccct ggagggcgcc cgcagacggt gctgaccgac gacgagatcg 120
aggagttcct cgagggcttc ctgtcagagc tagaacctga gccccgggag gacgacgtgg 180
aggccccgcc gcctcccgag cccaccccgc gggtccgaaa agcccaggcg gggggcaagc 240
cagggaagcg gccagggacg gccgcagaag tgcctccgga aaagaccaaa gacaaaggga 300
agaaaggcaa gaaagacaaa ggccccaagg tgcccaagga gtccttggag gggtccccca 360
ggccgcccaa gaaggggaag gagaagccac ccaaggccac caagaagccc aaggagaagc 420
cacctaaggc caccaagaag cccaaggaga agccacccaa ggccaccaag aagcccaaag 480
agaagccacc caaggccacc aagaagcccc cgtcagggaa gaggcccccc attctggctc 540
cctcagaaac cctggagtgg ccactgcccc caccccccag ccctggcccc gaggagctac 600
cccaggaggg aggggcgccc ctctcaaata actggcagaa tccaggagag gagacccatg 660
tggaggcacg ggagcaccag cctgagccgg aggaggagac cgagcaaccc acactggact 720
acaatgacca gatcgagagg gaggactatg aggactttga gtacattcgg cgccagaagc 780
aacccaggcc acccccaagc agaaggagga ggcccgagcg ggtctggcca gagacccctg 840
aggagaaggc cccggcccca gccccggagg agaggattga gcctcctgtg aagcctctgc 900
tgcccccgct gcccccctgac tatggtgatg gttacgtgat ccccaactac gatgacatgg 960
actattactt tgggcctcct ccgccccaga agcccgatgc tgagcgccag acggacgaag 1020
agaaggagga gctgaagaaa cccaaaaagg aggacagcag ccccaaggag gagaccgaca 1080
agtgggcagt ggagaagggc aaggaccaca aagagccccg aaagggcgag gagttggagg 1140
aggagtggac gcctacggag aaagtcaagt gtccccccat tgggatggag tcacaccgta 1200
ttgaggacaa ccagatccga gcctcctcca tgctgcgcca cggcctgggg gcacagcgcg 1260
gccggctcaa catgcagacc ggtgccactg aggacgacta ctatgatggt gcgtggtgtg 1320
ccgaggacga tgccaggacc cagtggatag aggtggacac caggaggact acccggttca 1380
caggcgtcat cacccagggc agagactcca gcatccatga cgattttgtg accaccttct 1440

tcgtgggctt cagcaatgac agccagacat gggtgatgta caccaacggc tatgaggaaa 1500
tgacctttca tgggaacgtg gacaaggaca cacccgtgct gagtgagctc ccagagccgg 1560
tggtggctcg tttcatccgc atctacccac tcacctggaa tggcagcctg tgcatgcgcc 1620
tggaggtgct ggggtgctct gtggcccctg tctacagcta ctacgcacag aatgaggtgg 1680
tggccaccga tgacctggat ttccggcacc acagctacaa ggacatgcgc cagctcatga 1740
aggtggtgaa cgaggagtgc cccaccatca cccgcactta cagcctgggc aagagctcac 1800
gaggcctcaa gatctatgcc atggagatct cagacaaccc tggggagcat gaactggggg 1860
agcccgagtt ccgctacact gctgggatcc atggcaacga ggtgctgggc cgagagctgt 1920
tgctgctgct catgcagtac ctgtgccgag agtaccgcga tgggaaccca cgtgtgcgca 1980
gcctggtgca ggacacacgc atccacctgg tgccctcact gaaccctgat ggctacgagg 2040
tggcagcgca gatgggctca gagtttggga actgggcgct gggactgtgg actgaggagg 2100
gctttgacat ctttgaagat ttcccggatc tcaactctgt gctctgggga gctgaggaga 2160
ggaaatgggt cccctaccgg gtccccaaca ataacttgcc catccctgaa cgctaccttt 2220
cgccagatgc cacggtatcc acggaggtcc gggccatcat tgcctggatg gagaagaacc 2280
ccttcgtgct gggagcaaat ctgaacggcg gcgagcggct agtatcctac ccctacgata 2340
tggcccgcac gcctacccag gagcagctgc tggccgcagc catggcagca gcccgggggg 2400
aggatgagga cgaggtctcc gaggcccagg agactccaga ccacgccatc ttccggtggc 2460
ttgccatctc cttcgcctcc gcacacctca ccttgaccga gccctaccgc ggaggctgcc 2520
aagcccagga ctacaccggc ggcatgggca tcgtcaacgg ggccaagtgg aacccccgga 2580
ccgggactat caatgacttc agttacctgc ataccaactg cctggagctc tccttctacc 2640
tgggctgtga caagttccct catgagagtg agctgccccg cgagtgggag aacaacaagg 2700
aggcgctgct caccttcatg gagcaggtgc accgtggcat taatcacggc gtgaagacag 2760
ccagtggtgg tgattactgg cgaatcttga acccgggtga gtaccgcgtg acagcccacg 2820
cggagggcta caccccgagc gccaagacct gcaatgttga ctatgacatc ggggccactc 2880
agtgcaactt catcctggct cgctccaact ggaagcgcat ccgggagatc atggccatga 2940
acgggaaccg gcctatccca cacatagacc catcgcgccc tatgaccccc caacagcgac 3000
gcctgcagca gcgacgccta caacaccgcc tgcggcttcg ggcacagatg cggctgcggc 3060
gcctcaacgc caccaccacc ctaggccccc acactgtgcc tcccacgctg ccccctgccc 3120
ctgccaccac cctgagcact accatagagc cctggggcct cataccgcca accaccgctg 3180
gctgggagga gtcggagact gagacctaca cagaggtggt gacagagttt gggaccgagg 3240
tggagcccga gtttgggacc aaggtggagc ccgagtttga gacccagttg gagcctgagt 3300
ttgagaccca gctggaaccc gagtttgagg aagaggagga ggaggagaaa gaggaggaga 3360
tagccactgg ccaggcattc cccttcacaa cagtagagac ctacacagtg aactttgggg 3420
acttctgaga tcagcgtcct accaagaccc cagcccaact caagctacag cagcagcact 3480
tcccaagcct gctgaccaca gtcacatcac ccatcagcac atggaaggcc cctggtatgg 3540
acactgaaag gaagggctgg tcctgcccct ttgagggggt gcaaacatga ctgggaccta 3600
agagccagag gctgtgtaga ggctcctgct ccacctgcca gtctcgtaag agatggggtt 3660
gctgcagtgt tggagtaggg gcagagggag ggagccaagg tcactccaat aaaacaagct 3720
catggcaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3780
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaagaaggg 3840
ggggcggcgc gtgggtgggc ccccgcgcgg gg 3872

<210> 102
<211> 2545
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510010CB1

<400> 102
cggctcgagg tgagtttggc cactccgtag tgtgcacttg gtgagggcag cagctcgcca 60
cagctgccag ccatctgtcc attcacccat ctgtccatct ggcagcccgc tgttcagacc 120
tgtctgtctg tccgcccatc tgtaagccca tctctgtccc attgtctatc tgaccatctt 180
tctcttactg tcctctttgt ctagctatct ggcctatctg tcgatccatc ttcgtgtctg 240
tcttcagccc ccacctgttt gtccatctgt ccaattacct gtgactctgt gcatcttctt 300
gtccattcat ctgcccaccc atccgtccct ccgtctgccc accagccgcc cctctcctcc 360
tgggctgcag agccatggcc cggggctacg ggccacggt cagcctagtc ctgctgggtc 420
tggggctggc gctggctgtc attgtgctgg ctgtggtcct ctctcgacac caggccccat 480
gtggccccca ggcctttgcc cacgctgctg ttgccgccga ctccaaggtc tgctcggata 540
ttggacgagc catcctccag cagcagggct cacccgtgga tgccaccatc gcggctctgg 600
tctgcaccag cgtcgtcaac cctcagagca tgggcctggg cggaggggtc atcttcacca 660

tctacaatgt gacaacaggg aaggtggagg tcatcaatgc ccgggagacg gtgccggcca 720
gccacgcccc gagcctgctg gaccagtgtg cacaggctct gccactgggc acaggggccc 780
agtggatcgg ggtgcccggg gagctccgtg gctatgccga ggcccaccgc cgtcatggcc 840
gcctgccctg ggcgcagctg ttccagccca ccatcgcgct gctccgaggg gggcatgtgg 900
tggcccctgt cctcagccgt ttcctgcaca acagcatcct gcggccttcc ttgcaggcgt 960
caaccctgcg ccagctcttc ttcaacggga cagaacccct gaggcctcag gacccactcc 1020
catggcctgc actggccacc accctggaga ccgtggccac agagggcgtg gaggtcttct 1080
acacggggag gctgggccag atgctggtgg aggacattgc caaggaaggg agccagctga 1140
cgctgcagga cctggccaag ttccagcccg aggtggtgga tgccctggag gtgcccctgg 1200
gggactatac cctgtactca ccaccgccgc ctgcaggggg tgccattctc agctttatcc 1260
tcaacgtgct aagagggttc aacttctcaa cagagtctat ggccaggcct gaagggaggg 1320
tgaacgtgta ccaccacctt gtagagacgc tcaagtttgc caggggggcag aggtggaggc 1380
tgggggaccc tcgaagccac ccgaagctcc agaatgcctc ccgggacctg ctgggggaga 1440
ccctggccca gctcatccgc caacagatcg atggccgggg ggaccaccag ctcagccact 1500
acagcttggc cgaggcctgg ggccacggga caggcacgtc ccatgtgtct gtgctggggg 1560
aggatggcag cgccgtggct gccaccagca ccatcaacac accctttgga gcgatggtgt 1620
attcaccacg gacaggcatc atcctcaaca acgagctcct ggacttatgc gagcgatgcc 1680
cccggggttc cggcaccacc ccctcacctg tgagtggaga cagggtgggt ggagctcccg 1740
gaaggtgctg gcccccagtt ccaggcgagc gttccccatc ctccatggtg ccctccatct 1800
tgatcaacaa agcccagggg tcgaagctag tgattggcgg ggctggcggg gagctcatca 1860
tctctgctgt ggcccaggcc atcatgagca agctgtggct tggctttgac ctgagagcgg 1920
ccattgcagc ccccatcctg catgtcaaca gcaagggctg tgtggagtac gagcccaact 1980
tcagccagaa acaggatctt gctgcattgc ccaggctggt ctcaaactcc tggcctcaag 2040
tgatcctcct tcctcagcct cccaaattgc tggcattaca ggaggtgcag aggggactcc 2100
aagaccgtgg ccagaaccag acccagaggc ccttcttcct gaacgtggtc caggctgtgt 2160
cccaggaggg ggcctgtgtg tacgccgtct cggacctgag gaagagtggg gaggccgcag 2220
gctactaaga cactgctctg cccagagctg aagtctggcc ccaccatgag tcctgtgtcc 2280
aggccggaca tggctggggg accaactact ctggcaggat ctggacccct ggcaggggag 2340
tccagctgag agtggaagag gtggcgggga ccagctgggc agatgagagg ctgagcctca 2400
tccctaaccc cctttcccag agcccctggt ggtcctgaac cggcccctct atccctccgc 2460
aggcctcttg cctggggcca ctctcccacc ctctcgatct gtatatcctc cagtccaaga 2520
ttaaagaggc ggactgtgaa aaaaa 2545

<210> 103
<211> 3273
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510056CB1

<400> 103
gaggactatg aggactgtga gtagggtcct gccagcccca cctgggtcgg accccctggcc 60
tgggggatgt gccaatgggc ccatcccagc cttgggcccc actctgagcc agcctccccc 120
tcagttgagt acattcggcg ccagaagcaa cccaggccac ccccaagcag aaggaggagg 180
cccgagcggg tctggccaga cccccctgag gagaaggccc cggccccagc cccggaggag 240
aggattgagc ctcctgtgaa gcctctgctg cccccgctgc cccctgacta tggtgatggt 300
tacgtgatcc ccaactacga tgacatggac tattactttg ggcctcctcc gccccagaag 360
cccgatgctg agcgccagac ggacgaagag aaggaggagc tgaagaaacc caaaaaggag 420
gacagcagcc ccaaggagga gaccgacaag tgggcagtgg agaagggcaa ggaccacaaa 480
gagccccgaa agggcgagga gttggaggag gagtggacgc ctacggagaa agtcaagtgt 540
ccccccattg ggatggagtc acaccgtatt gaggacaacc agatccgagc ctcctccatg 600
ctgcgccacg gcctgggggc acagcgcggc cggctcaaca tgcagaccgg tgccactgag 660
gacgactact atgatggtgc gtggtgtgcc gaggacgatg ccaggaccca gtggatagag 720
gtggacacca ggaggactac ccggttcaca ggcgtcatca cccagggcag agactccagc 780
atccatgacg attttgtgac caccttcttc gtgggcttca gcaatgacag ccagacatgg 840
gtgatgtaca ccaacggcta tgaggaaatg gtgggcacca tgcccaggct cttggctctg 900
ctcccattgt ctgggcgagg ggtgggctct cagaggggct ggcagtactg ctctgaggcc 960
tgcctctccc cagacctttc atgggaacgt ggacaaggac acacccgtgc tgagtgagct 1020
cccagagccg gtggtggctc gtttcatccg catctaccca ctcacctgga atggcagcct 1080
gtgcatgcgc ctggaggtgc tggggtgctc tgtggcccct gtctacagct actacgcaca 1140
gaatgaggtg gtggccaccg atgacctgga tttccggcac cacagctaca aggacatgcg 1200

ccagctcatg aaggtggtga acgaggagtg ccccaccatc acccgcactt acagcctggg 1260
caagagctca cgaggcctca agatctatgc catggagatc tcagacaacc ctggggagca 1320
tgaactgggg gagcccgagt tccgctacac tgctgggatc catggcaacg aggtgctggg 1380
ccgagagctg ttgctgctgc tcatgcagta cctgtgccga gagtaccgcg atgggaaccc 1440
acgtgtgcgc agcctggtgc aggacacacg catccacctg gtgccctcac tgaaccctga 1500
tggctacgag gtggcagcgc agatgggctc agagtttggg aactgggcgc tgggactgtg 1560
gactgaggag ggctttgaca tctttgaaga tttcccggat ctcaactctg tgctctgggg 1620
agctgaggag aggaaatggg tcccctaccg ggtccccaac aataacttgc ccatccctga 1680
acgctacctt tcgccagatg ccacggtatc cacggaggtc cgggccatca ttgcctggat 1740
ggagaagaac cccttcgtgc tgggagcaaa tctgaacggc ggcgagcggc tagtatccta 1800
cccctacgat atggcccgca cgcctaccca ggagcagctg ctggccgcag ccatggcagc 1860
agcccggggg gaggatgagg acgaggtctc cgaggcccag gagactccag accacgccat 1920
cttccggtgg cttgccatct ccttcgcctc cgcacacctc accttgaccg agccctaccg 1980
cggaggctgc caagcccagg actacaccgg cggcatgggc atcgtcaacg gggccaagtg 2040
gaacccccgg accgggacta tcaatgactt cagttacctg cataccaact gcctggagct 2100
ctccttctac ctgggctgtg acaagttccc tcatgagagt gagctgcccc gcgagtggga 2160
gaacaacaag gaggcgctgc tcaccttcat ggagcaggtg caccgcggca ttaagggggt 2220
ggtgacggac gagcaaggca tccccattgc caacgccacc atctctgtga gtggcattaa 2280
tcacggcgtg aagacagcca gtggtggtga ttactggcga atcttgaacc cgggtgagta 2340
ccgcgtgaca gcccacgcgg agggctacac cccgagcgcc aagacctgca atgttgacta 2400
tgacatcggg gccactcagt gcaacttcat cctggctcgc tccaactgga agcgcatccg 2460
ggagatcatg gccatgaacg ggaaccggcc tatcccacac atagacccat cgcgccctat 2520
gaccccccaa cagcgacgcc tgcagcagcg acgcctacaa caccgcctgc ggcttcgggc 2580
acagatgcgg ctgcggcgcc tcaacgccac caccacccta ggcccccaca ctgtgcctcc 2640
cacgctgccc cctgcccctg ccaccaccct gagcactacc atagagccct ggggcctcat 2700
accgccaacc accgctggct gggaggagtc ggagactgag acctacacag aggtggtgac 2760
agagtttggg accgaggtgg agcccgagtt tgggaccaag gtggagcccg agtttgagac 2820
ccagttggag cctgagtttg agacccagct ggaacccgag tttgaggaag aggaggagga 2880
ggagaaagag gaggagatag ccactggcca ggcattcccc ttcacaacag tagagaccta 2940
cacagtgaac tttggggact tctgagatca gcgtcctacc aagaccccag cccaactcaa 3000
gctacagcag cagcacttcc caagcctgct gaccacagtc acatcaccca tcagcacatg 3060
gaaggcccct ggtatggaca ctgaaaggaa gggctggtcc tgcccctttg agggggtgca 3120
aacatgactg ggacctaaga gccagaggct gtgtagaggc tcctgctcca cctgccagtc 3180
tcgtaagaga tggggttgct gcagtgttgg agtaggggca gagggaggga gccaaggtca 3240
ctccaataaa acaagctcat ggcaaaaaaa aaa 3273

<210> 104
<211> 1409
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510398CB1

<400> 104
gccagggacc aagatggatc ttctcctcga catcagctaa gcctggagga ctcttcccct 60
cagagaccat ggagagggac agccacggga atgcatctcc agcaagaaca ccttcagctg 120
gagcatctcc agcccaggca tctccagctg ggacacctcc aggccgggca tctccagccc 180
aggcatctcc agcccaggca tctccagctg ggacacctcc gggccgggca tctccagccc 240
aggcatctcc agctggtaca cctccaggcc gggcatctcc aggccgggca tctccagccc 300
aggcatctcc agcccgggca tctccggctc tggcatcact ttccaggtcc tcatccggca 360
ggtcatcatc cgccaggtca gcctcggtga caacctcccc aaccagagtg taccttgtta 420
gagcaacacc agtggggcct gtacccatcc gatcatctcc tgccaggtca gcaccagcaa 480
ccagggccac cagggagagc ccaggtacga gcctgcccaa gttcacctgg cgggagggcc 540
agaagcagct accgctcatc gggtgcgtgc tcctcctcat tgccctggtg gtttcgctca 600
tcatcctctt ccagttctgg cagggccaca cagggatcag gtacaaggag cagagggaga 660
gctgtcccaa gcacgctgtt cgctgtgacg ggtggtagaa ctgcaagctg aagagtgacg 720
agctgggctg cgtgaggttt gactgggaca agtctctgct taaaatctac tctgggtcct 780
cccatcagtg gcttcccatc tgtagcagca actggaatga ctcctactca gagaagacct 840
gccagcagct gggtttcgag agtgctcacc ggacaaccga ggttgcccac agggattttg 900
ccaacagctt ctcaatcttg agatacaact ccaccatcca ggaaagcctc cacaggtctg 960
aatgcccttc ccagcggtat atctccctcc agtgttccca ctgcggactg agggccatga 1020

ccgggcggat cgtgggaggg gcgctggcct cggatagcaa gtggccttgg caagtgagtc 1080
tgcacttcgg caccacccac atctgtggag gcacgctcat tgacgcccag tgggtgctca 1140
ctgccgccca ctgcttcttc gtgacccggg agaaggtcct ggagggctgg aaggtgtacg 1200
cgggcaccag caactcatat ccgggcccca aagcttctgc agggcagaag tcaaagactc 1260
ttaaagatcc ttacatggaa cacttctgtt ttataattag ggaaactgaa gcccaagggt 1320
tataaataag tttgctccaa atgacacatc agctaagcct ctcctcgaca tcagctaagc 1380
ctgaataagt ttgctccaaa tgacacatc 1409

<210> 105
<211> 4025
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510498CB1

<400> 105
tctcccgccc cttcctggat tccctcaccc gtctcgatcc cctctccgcc ctttcccaga 60
gacccagagc ccctgacccc ccgcgccctc cccggagccc cccgcgcgtg ccgcggccat 120
ggcggccgtg cgcggggcgc ccctgctcag ctgcctcctg gcgttgctgg ccctgtgccc 180
tggagggcgc ccgcagacgg tgctgaccga cgacgagatc gaggagttcc tcgagggctt 240
cctgtcagag ctagaacctg agccccggga ggacgacgtg gaggccccgc cgcctcccga 300
gcccaccccg cgggtccgaa aagcccaggc gggggggcaag ccagggaagc ggccagggac 360
ggccgcagaa gtgcctccgg aaaagaccaa agacaaaggg aagaaaggca agaaagacaa 420
aggccccaag gtgcccaagg agtccttgga ggggtccccc aggccgccca agaaggggaa 480
ggagaagcca cccaaggcca ccaagaagcc caaggagaag ccacctaagg ccaccaagaa 540
gcccaaggag aagccaccca aggccaccaa gaagcccaaa gagaagccac ccaaggccac 600
caagaagccc ccgtcaggga agaggccccc cattctggct ccctcagaaa ccctggagtg 660
gccactgccc ccacccccca gccctggccc cgaggagcta ccccaggagg gaggggcgcc 720
cctctcaaat aactggcaga atccaggaga ggagacccat gtggaggcac gggagcaccg 780
gcctgagccg gaggaggaga ccgagcaacc cacactggac tacaatgacc agatcgagag 840
ggaggactat gaggactttg agtacattcg gcgccagaag caacccaggc cacccccaag 900
cagaaggagg aggcccgagc gggtctggcc agagacccct gaggagaagg ccccggcccc 960
agccccggag gagaggattg agcctcctgt gaagcctctg ctgcccccgc tgcccccctga 1020
ctatggtgat ggttacgtga tccccaacta cgatgacatg gactattact ttgggcctcc 1080
tccgccccag aagcccgatg ctgagcgcca gacggacgaa gagaaggagg agctgaagaa 1140
acccaaaaag gaggacagca gccccaagga ggagaccgac aagtgggcag tggagaaggg 1200
caaggaccac aaagagcccc gaaagggcga ggagttggag gaggagtgga cgcctacgga 1260
gaaagtcaag tgtcccccca ttgggatgga gtcacaccgt attgaggaca accagatccg 1320
agcctcctcc atgctgcgcc acggcctggg ggcacagcgc ggccggctca acatgcagac 1380
cggtgccact gaggacgact actatgatgg tgcgtggtgt gccgaggacg atgccaggac 1440
ccagtggata gaggtggaca ccaggaggac tacccggttc acaggcgtca tcacccaggg 1500
cagagactcc agcatccatg acgattttgt gaccaccttc ttcgtgggct tcagcaatga 1560
cagccagaca tgggtgatgt acaccaacgg ctatgaggaa atggtgggca ccatgcccag 1620
gctcttggct ctgctcccat tgtctgggcg aggggtgggc tctcagaggg gctggcagta 1680
ctgctctgag gcctgcctct ccccagacct ttcatgggaa cgtggacaag gacacacccg 1740
tgctgagtga gctcccagag ccggtggtgg ctcgtttcat ccgcatctac ccactcacct 1800
ggaatggcag cctgtgcatg cgcctggagg tgctggggtg ctctgtggcc cctgtctaca 1860
gctactacgc acagaatgag gtggtggcca ccgatgacct ggatttccgg caccacagct 1920
acaaggacat gcgccagctc atgaaggtgg tgaacgagga gtgccccacc atcacccgca 1980
cttacagcct gggcaagagc tcacgaggcc tcaagatcta tgccatggag atctcagaca 2040
accctgggga gcatgaactg gggggagcccg agttccgcta cactgctggg atccatggca 2100
acgaggtgct gggccgagag ctgttgctgc tgctcatgca gtacctgtgc cgagagtacc 2160
gcgatgggaa cccacgtgtg cgcagcctgg tgcaggacac acgcatccac ctggtgccct 2220
cactgaaccc tgatggctac gaggtggcag cgcagatggg ctcagagttt gggaactggg 2280
cgctgggact gtggactgag gagggctttg acatctttga agatttcccg gatctcaact 2340
ctgtgctctg ggagctgag gagaggaaat gggtccccta ccgggtcccc aacaataact 2400
tgcccatccc tgaacgctac ctttcgccag atgccacggt atccacggag gtccgggcca 2460
tcattgcctg gatggagaag aaccccttcg tgctgggagc aaatctgaac ggcggcgagc 2520
ggctagtatc ctaccccctac gatatggccc gcacgcctac ccaggagcag ctgctggccg 2580
cagccatggc agcagcccgg ggggaggatg aggacgaggt ctccgaggcc caggagactc 2640
cagaccacgc catcttccgg tggcttgcca tctccttcgc ctccgcacac ctcaccttga 2700

ccgagcccta ccgcggaggc tgccaagccc aggactacac cggcggcatg ggcatcgtca 2760
acggggccaa gtggaacccc cggaccggga ctatcaatga cttcagttac ctgcatacca 2820
actgcctgga gctctccttc tacctgggct gtgacaagtt ccctcatgag agtgagctgc 2880
cccgcgagtg ggagaacaac aaggaggcgc tgctcacctt catggagcag gtgcaccgcg 2940
gcattaaggg ggtggtgacg gacgagcaag gcatccccat tgccaacgcc accatctctg 3000
tgagtggcat taatcacggc gtgaagacag ccagtggtgg tgattactgg cgaatcttga 3060
acccgggtga gtaccgcgtg acagcccacg cggagggcta cacccсgagc gccaagacct 3120
gcaatgttga ctatgacatc ggggccactc agtgcaactt catcctggct cgctccaact 3180
ggaagcgcat ccgggagatc atggccatga acgggaaccg gcctatccca cacatagacc 3240
catcgcgccc tatgaccccc caacagcgac gcctgcagca gcgacgccta caacaccgcc 3300
tgcggcttcg ggcacagatg cggctgcggc gcctcaacgc caccaccacc ctaggccccc 3360
acactgtgcc tcccacgctg cccccctgccc ctgccaccac cctgagcact accatagagc 3420
cctggggcct cataccgcca accaccgctg gctgggagga gtcggagact gagacctaca 3480
cagaggtggt gacagagttt gggaccgagg tggagcccga gtttgggacc aaggtggagc 3540
ccgagtttga gacccagttg gagcctgagt ttgagaccca gctggaaccc gagtttgagg 3600
aagaggagga ggaggagaaa gaggaggaga tagccactgg ccaggcattc cccttcacaa 3660
cagtagagac ctacacagtg aactttgggg acttctgaga tcagcgtcct accaagaccc 3720
cagcccaact caagctacag cagcagcact tcccaagcct gctgaccaca gtcacatcac 3780
ccatcagcac atggaaggcc cctggtatgg acactgaaag gaagggctgg tcctgcccct 3840
ttgagggggt gcaaacatga ctgggaccta agagccagag gctgtgtaga ggctcctgct 3900
ccacctgcca gtctcgtaag agatggggtt gctgcagtgt tggagtaggg gcagagggag 3960
ggagccaagg tcactccaat aaaacaagct catggcacgg acaaagaaaa gatgacaaaa 4020
aaacg 4025

<210> 106
<211> 2951
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510044CB1

<400> 106
gatggctacg gcagggggtg gctctggggc tgacccggga agtcggggtc tccttcgcct 60
tctgtctttc tgcgtcctac tagcaggcct cttcctggac ttcgagcctg accctctccc 120
actttgtcca gatgtctcgt tttcccacaa accccagcca cccaccccac agtcgaaagg 180
aacaggtttg tgcaggggaa actcagtgga gaggaagata tatcccccct taaataaaac 240
agctccctgt gttcgcctgc tcaacgccac tcatcagatt ggctgccagt cttcaattag 300
tggagacaca ggggttatcc acgtagtaga gaaagaggag gacctacagt gggtattgac 360
tgatggcccc aacccccctt acatggttct gctggagagc aagcatttta ccagggattt 420
aatggagaag ctgaaaggga gaaccagccg aattgctggt cttgcagtgt ccttgaccaa 480
gcccagtcct gcctcaggct tctctcctag tgtacagtgc ccaaatgatg ggtttggtgt 540
ttactccaat tcctatgggc cagagtttgc tcactgcaga gaaatacagt ggaattcgct 600
gggcaatggt ttggcttatg aagactttag tttccccatc tttcttcttg aagatgaaaa 660
tgaaaccaaa gtcatcaagc agtgctatca agatcacaac ctgagtcaga atggctcagc 720
accaaccttc ccactatgtg ccatgcagct cttttcacac atgcatgctg tcatcagcac 780
tgccacctgc atgcggcgca ggtccatcca aagcaccttc agcatcaacc cagaaatcgt 840
ctgtgacccc ctgtctgatt acaatgtgtg gagcatgcta aagcctataa atacaactgg 900
gacattaaag cctgacgaca gggttgtggt tgctgccacc cggctggata gtcgttcctt 960
tttctggaat gtggccccag gggctgaaag cgcagtggct tcctttgtca cccagctggc 1020
tgctgctgaa gctttgcaaa aggcacctga tgtgaccacc ctgccccgca atgtcatgtt 1080
tgtcttcttt caaggggaaa cttttgacta cattggcagc tcgaggatgg tctacgatat 1140
ggagaagggc aagtttcccg tgcagttaga gaatgttgac tcatttgtgg agctgggaca 1200
ggtggcctta agaacttcat tagagctttg gatgcacaca gatcctgttt ctcagaaaaa 1260
tgagtctgta cggaaccagg tggaggatct cctggccaca ttggagaaga gtggtgctgg 1320
tgtccctgct gtcatcctca ggaggccaaa tcagtcccag cctctcccac catcttccct 1380
gcagcgattt cttcgagctc gaaacatctc tggcgttgtt ctggctgacc actctggtgc 1440
cttccataac aaatattacc agagtattta cgacactgct gagaacatta atgtgagcta 1500
tcccgaatgg ctgagccctg aagaggacct gaactttgta acagacactg ccaaggccct 1560
ggcagatgtg gccacggtgc tgggacgtgc tctgtatgag cttgcaggag gaaccaactt 1620
cagcgacaca gttcaggctg atccccaaac ggttacccgc ctgctctatg ggttcctgat 1680
taaagccaac aactcatggt tccagtctat cctcaggcag gacctaaggt cctacttggg 1740

tgacgggcct cttcaacatt acatcgctgt ctccagcccc accaacacca cttatgttgt 1800
acagtatgcc ttggcaaatt tgactggcac agtggtcaac ctcacccgag agcagtgcca 1860
ggatccaagt aaagtcccaa gtgaaaacaa ggatctgtat gagtactcat gggtccaggg 1920
ccctttgcat tctaatgaga cggaccgact cccccggtgt gtgcgttcta ctgcacgatt 1980
agccagggcc ttgtctcctg cctttgaact gagtcagtgg agctctactg aatactctac 2040
atggactgag agccgctgga aagatatccg tgcccggata tttctcatcg ccagcaaaga 2100
gcttgagttg atcaccctga cagtgggctt cggcatcctc atcttctccc tcatcgtcac 2160
ctactgcatc aatgccaaag ctgatgtcct tttcattgct ccccgggagc caggagctgt 2220
gtcatactga ggaggacccc agcttttctt gccagctcag cagttcactt cctagagcat 2280
ctgtcccact gggacacaac cactaatttg tcactggaac ctccctgggc ctgtctcaga 2340
ttgggattaa cataaaagag tggaactatc caaaagagac agggagaaat aaataaattg 2400
cctcccttcc tccgctcccc tttcccatca cccctttccc atttcctctt ccttctctac 2460
tcatgccaga ttttgggatt acaaatagaa gcttcttgct cctgtttaac tccctagtta 2520
cccaccctaa tttgcccttc aggacccttc tactttttcc ttcctgccct gtacctctct 2580
ctgctcctca cccccacccc tgtacccagc caccttcctg actgggaagg acataaaagg 2640
tttaatgtca gggtcaaact acattgagcc cctgaggaca ggggcatctc tgggctgagc 2700
ctactgtctc cttcccactg tcctttctcc aggccctcag atggcacatt agggtgggcg 2760
tgctgcgggt gggtatccca cctccagccc acagtgctca gttgtacttt ttattaagct 2820
gtaatatcta tttttgtttt tgtctttttc ctttattctt tttgtaaata tatatataat 2880
gagtttcatt aaaatagatt atcccacaaa aaaaaaaaaa aagggcggcc gctcgcgatc 2940
tagaactagt c 2951

<210> 107
<211> 1832
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504509CB1

<400> 107
cccacgcgtc cgggcggcgc ggccggggaa gcgtgatgaa ggcctacgag tgcggcgcgg 60
cctgaagggg cacgcggggg acctgcaaag ctagtgaggg gcggggcagg cggcgcggtg 120
ggggcgggcc gagcccggag gccagatgag cggacacagc cccacgcgcg gggccatgca 180
ggtggccatg aacggtaagg cccgcaaaga ggcggtgcag actgcggcta aggaactcct 240
caagttcgtg aaccggagtc cctctccttt ccatggctgg cttcagtgaa ctcaaggaga 300
ctgagaaatg gaatattaag cccgagagca agtacttcat gaccaggaac tcctccacca 360
tcatagcttt tgctgtaggg ggccagtacg ttcctggcaa tggcttcagc ctcatcgggg 420
cccacacgga cagcccctgc ctccgggtga aacgtcggtc tcgccgcagc caggtgggct 480
tccagcaagt cggtgtggag acctatggtg gtgggatctg gagcacctgg tttgaccgtg 540
acctgactct ggctggacgc gtcattgtca agtgccctac ctcaggtcgg ctggagcagc 600
agctggtgca cgtggagcgg cccattcttc gcatcccaca cctggccatc catctgcagc 660
gaaatatcaa cgagaacttt gggcccaaca cagagatgca tctagtcccc attcttgcca 720
cagccatcca ggaggagctg gagaaggggа ctcctgagcc agggcctctc aatgctgtgg 780
atgagcggca ccattcggtc ctcatgtccc tgctctgtgc ccatctgggg ctgagcccca 840
aggacatagt ggagatggag ctctgccttg cagacaccca gcctgcggtc ttgggtggtg 900
cctatgatga gttcatcttt gctcctcggc tggacaatct gcacagctgc ttctgtgccc 960
tgcaggcctt gatagattcc tgtgcaggcc ctggctccct ggccacagag cctcacgtgc 1020
gcatggtcac actctatgac aacgaagagg tggggtctga gagtgcacag ggagcacagt 1080
cactgctgac agagctggtg ctgcggcgga tctcagcctc gtgccagcac ccgacagcct 1140
tcgaggaagc catacccaag tccttcatga tcagcgcaga catggcccat gctgtgcatc 1200
ccaactacct ggacaagcat gaggagaacc accggccttt attccacaag ggccccgtga 1260
tcaaggtgaa cagcaagcaa cgctatgctt caaacgcggt gtcagaggcc ctgatccgag 1320
aggtggccaa caaagtcaag gtccccctgc aggatctcat ggtccggaat gacacccccт 1380
gtggaaccac cattggacct atcttggctt ctcggctggg gctgcgggtg ctggatttag 1440
gcagccccca actggccatg cactctatcc gggagatggc ctgcaccaca ggagtcctcc 1500
agaccctcac cctcttcaag ggcttctttg agctgttccc ttctctaagc cataatctct 1560
tagtggattg agccctcttg gaaagacttc tctgccatcc ctttgcacct gagaggggaa 1620
gttctcagct gagctgaagc tggattatta aagtggattg tcactcagac tctccgtgct 1680
acgcttattt ggagactaga ggagtgggag ttgagcctgg cttgaacctt tggaaccaga 1740
aaagttgggg agcaggtgga ggaggccaca ctcctgggag ctgatggttt taaatctggt 1800
tttaaatctc tgaaaaaaaa aaaaagaaaa aa 1832

<210> 108
<211> 1471
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506825CB1

<400> 108
cgacatccgg cctcggccac gtggtgggcg gaccggggcg gtcctgagcc tgcgacctcg 60
caggcgacct cgctggaccc taagtccagg ccacagtcag ggaagggcgc tgagaggcga 120
gcgtgagccc agcgacagga gagtgagctc accacgcgca gcgccatgac cagcaagggt 180
cccgaggagg agcacccatc ggtgacgctc ttccgccagt acctgcgtat ccgcactgtc 240
cagcccaagc ctgactatgg tggcacctgg ctatgtggtg accgtgttga cctggccagg 300
caccaaccct acactctcct ccatcttgct caactcccac acggatgtgg tgcctgtctt 360
caaggaacat tggagtcacg accccttga ggccttcaag gattctgagg gctacatcta 420
tgccaggggt gcccaggaca tgaagtgcgt cagcatccag tacctggaag ctgtgaggag 480
gctgaaggtg gagggccacc ggttccccag aaccatccac atgacctttg tgcctgatga 540
ggaggttggg ggtcaccaag gcatggagct gttcgtgcag cggcctgagt tccacgccct 600
gagggcaggc tttgccctgg atgagggcat agccaatccc actgatgcct tcactgtctt 660
ttatagtgag cggagtccct ggtgggtgcg ggttaccagc actgggaggc caggccatgc 720
ctcacgcttc atggaggaca cagcagcaga gaagctgcac aaggttgtaa actccatcct 780
ggcattccgg gagaaggaat ggcagaggct gcagtcaaac cccccacctga aagaggggtc 840
cgtgacctcc gtgaacctga ctaagctaga gggtggcgtg gcctataacg tgatacctgc 900
caccatgagc gccagctttg acttccgtgt ggcaccggat gtggacttca aggcttttga 960
ggagcagctg cagagctggt gccaggcagc tggcgagggg gtcaccctag agtttgctca 1020
gaagtggatg cacccccaag tgacacctac tgatgactca aacccttggt gggcagcttt 1080
tagccgggtc tgcaaggata tgaacctcac tctggagcct gagatcatgc ctgctgccac 1140
tgacaaccgc tatatccgcg cggtgggggt cccagctcta ggcttctcac ccatgaaccg 1200
cacacctgtg ctgctgcacg accacgatga acggctgcat gaggctgtgt tcctccgtgg 1260
ggtggacata tatacacgcc tgctgcctgc ccttgccagt gtgcctgccc tgcccagtga 1320
cagctgagcc ctggaactcc taaacctttg cccctggggc ttccatccca accagtgcca 1380
aggacctcct cttcccctt ccaaataata aagtctatgg acagggctgt ctctgaagta 1440
ctaacacaaa aaaaaaaaaa aaaaaggggc g 1471

<210> 109
<211> 1344
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506828CB1

<400> 109
cgacatccgg cctcggccac gtggtgggcg gaccggggcg gtcctgagcc tgcgacctcg 60
caggcgacct cgctggaccc taagtccagg ccacagtcag ggaagggcgc tgagaggcga 120
gcgtgagccc agcgacagga gagtgagctc accacgcgca gcgccatgac cagcaagggt 180
cccgaggagg agcacccatc ggtgacgctc ttccgccagt acctgcgtat ccgcactgtc 240
cagcccaagc ctgactatgg agctgctgtg gctttctttg aggagacagc ccgccagctg 300
ggcctgggct gtcagaaagt agaggtggca cctggctatg tggtgaccgt gttgacctgg 360
ccaggcacca accctacact ctcctccatc ttgctcaact cccacacgga tgtggtgcct 420
gtcttcaagg aacattggag tcacgacccc tttgaggcct tcaaggattc tgagggctac 480
atctatgcca ggggtgccca ggacatgaag tgcgtcagca tccagtacct ggaagctgtg 540
aggaggctga aggtggaggg ccaccggttc cccagaacca tccacatgac ctttgtgcct 600
gatgaggagg ttgggggtca ccaaggcatg gagctgttcg tgcagcggcc tgagttccac 660
gccctgaggg caggctttgc cctggatgag ggcatagcca atcccactga tgccttcact 720
gtcttttata gtgagcggag tccctggtgg gtgcgggtta ccagcactgg gaggccaggc 780
catgcctcac gcttcatgga ggacacagca gcagagaagc tggcttttga ggagcagctg 840
cagagctggt gccaggcagc tggcgagggg gtcaccctag agtttgctca gaagtggatg 900
cacccccaag tgacacctac tgatgactca aacccttggt gggcagcttt tagccgggtc 960
tgcaaggata tgaacctcac tctggagcct gagatcatgc ctgctgccac tgacaaccgc 1020

```
tatatccgcg cggtgggggt cccagctcta ggcttctcac ccatgaaccg cacacctgtg 1080
ctgctgcacg accacgatga acggctgcat gaggctgtgt tcctccgtgg ggtggacata 1140
tatacacgcc tgctgcctgc ccttgccagt gtgcctgccc tgcccagtga cagctgagcc 1200
ctggaactcc taaacctttg cccctggggc ttccatccca accagtgcca aggacctcct 1260
cttccccctt ccaaataata aagtctatgg acagggctgt ctctgaagta ctaacacaag 1320
gacaaaaaaa aaaaaaaaaa aaaa                                        1344

<210> 110
<211> 1429
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506831CB1

<400> 110
cgacatccgg cctcggccac gtggtgggcg gaccggggcg gtcctgagcc tgcgacctcg 60
caggcgacct cgctggaccc taagtccagg ccacagtcag ggaagggcgc tgagaggcga 120
gcgtgagccc agcgacagga gagtgagctc accacgcgca gcgccatgac cagcaagggt 180
cccgaggagg agcacccatc ggtgacgctc ttccgccagt acctgcgtat ccgcactgtc 240
cagcccaagc ctgactatgg caccaaccct acactctcct ccatcttgct caactcccac 300
acggatgtgg tgcctgtctt caaggaacat tggagtcacg accccttga ggccttcaag 360
gattctgagg gctacatcta tgccaggggt gcccaggaca tgaagtgcgt cagcatccag 420
tacctggaag ctgtgaggag gctgaaggtg gagggccacc ggttccccag aaccatccac 480
atgacctttg tgcctgatga ggaggttggg ggtcaccaag gcatggagct gttcgtgcag 540
cggcctgagt tccacgccct gagggcaggc tttgccctgg atgagggcat agccaatccc 600
actgatgcct tcactgtctt ttatagtgag cggagtccct ggtgggtgcg ggttaccagc 660
actgggaggc caggccatgc ctcacgcttc atggaggaca cagcagcaga gaagctgcac 720
aaggttgtaa actccatcct ggcattccgg gagaaggaat ggcagaggct gcagtcaaac 780
ccccacctga aagaggggtc cgtgacctcc gtgaacctga ctaagctaga gggtggcgtg 840
gcctataacg tgatacctgc caccatgagc gccagctttg acttccgtgt ggcaccggat 900
gtggacttca aggcttttga ggagcagctg cagagctggt gccaggcagc tggcgagggg 960
gtcaccctag agtttgctca gaagtggatg cacccccaag tgacacctac tgatgactca 1020
aacccttggt gggcagcttt tagccgggtc tgcaaggata tgaacctcac tctggagcct 1080
gagatcatgc ctgctgccac tgacaaccgc tatatccgcg cggtgggggt cccagctcta 1140
ggcttctcac ccatgaaccg cacacctgtg ctgctgcacg accacgatga acggctgcat 1200
gaggctgtgt tcctccgtgg ggtggacata tatacacgcc tgctgcctgc ccttgccagt 1260
gtgcctgccc tgcccagtga cagctgagcc ctggaactcc taaacctttg cccctggggc 1320
ttccatccca accagtgcca aggacctcct cttccccctt ccaaataata aagtctatgg 1380
acagggctgt ctctgaagta ctaacacaag gacaaaaaaa aaaaaaaaa             1429

<210> 111
<211> 2187
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509968CB1

<400> 111
gctgacgctc ctcgtgggcc cctcactcag gccccgcccg ccctgccctc ccctcccgat 60
cccggagcca tgtggcccct ggccctagtg atcgcctccc tgaccttggc cttgtcagga 120
ggtgtctccc aggagtcttc caaggttctc aacaccaatg ggaccagtgg gtttctccca 180
ggtggctaca cctgcttccc ccactctcag ccctggcagg ctgccctact agtgcaaggg 240
cggctactct gtgggggagt cctggtccac cccaaatggg tcctcactgc cgcacactgt 300
ctaaaggagg ggctcaaagt ttacctaggc aagcacgccc tagggcgtgt ggaagctggt 360
gagcaggtga gggaagttgt ccactctatc cccccaccctg aataccggag aagccccacc 420
cacctgaacc acgaccatga catcatgctt ctggagctgc agtccccggt ccagctcaca 480
ggctacatcc aaaccctgcc cctttcccac aacaaccgcc taacccctgg caccacctgt 540
cgggtgtctg gctggggcac caccaccagc cccccagggta tgcacccaca caggtggcct 600
gaggccccat aggagtggct ggggaaacag gggcagagat gggagggaag gtctgagtga 660
```

attaccccaa aactctacaa tgtgccaaca tccaacttcg ctcagatgag gagtgtcgtc 720
aagtctaccc aggaaagatc actgacaaca tgttgtgtgc cggcacaaaa gagggtggca 780
aagactcctg tgagggtgac tctgggggcc ccctggtctg taacagaaca ctgtatggca 840
tcgtctcctg ggagacttc ccatgtgggc aacctgaccg gcctggtgtc tacacccgtg 900
tctcaagata cgtcctgtgg atccgtgaaa caatccgaaa atatgaaacc cagcagcaaa 960
aatggttgaa gggcccacaa taaaagttga gaaatgtacc ggcttccatc ctgtcaccat 1020
gacttcctca catggtctgc ttagcccttc tctgctcctt attcccagtg ttccatttga 1080
accagtgatc catgtcctga aaaatgctca atctcagcta acattccatg tttcagaagc 1140
attcaggcac tgccaggctt gcagtctccc agatgttgca tccctgaaac atctcaacaa 1200
cctgaatgtc ccaacccaga caatggccca ggtctctcaa cttcatcagt gtggcttcta 1260
tgagcccaga tcaccacctg aacgttctgt ctgtggcaca ttcttaaata tttccatcag 1320
cccatctcaa caatatatgt cctataaatg gaccatcctt gacaacatcc tctaactctt 1380
caagtattta ttcaatgcca gtatcctaga ccttctattt tttgcactca agaaggctct 1440
agactcccat gatagttcat cctgaaaata ttctcttatg cccacaatct tctgccctga 1500
caacattctg tgtacctctg tgactcacca cagctaacat tggatcctca gaatatttca 1560
ttctcacact gttatgggtg tctcagaagt cccaacccaa cctacatccc acattcttcc 1620
aataccccac ctctgccaac attccctctc tgaatcaatg gcaccctagt ctctagagtt 1680
atagggttca gtataccaaa gggtcttctt gcctgaactt tattgtctac caaatattcc 1740
gtcttgtatc cctccatga acatccttgg tcagtgtccc ttgctgttac atctttgtgc 1800
atgaccctaa aatgtagtgc aaatccttgc tttggacaag ttataaaact cacagtctct 1860
gtgctttctc atctgtaaaa tgggttcata attttttta attgtaacat tattacaaga 1920
ataaatgtca agcatttatc actaaaaaaa aaaaaaaaa aaaaaaaaaa aaaaaaaaa 1980
aaaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaacaca aaacaggggg ggggcggcgg 2040
cgctagaaat gagacaccct agccccgcg aggacataat atcgccggca gcgagcacca 2100
ctagagaggg ggcagaggga acactataca ccccgcatgt ataacgcgca cacacagagg 2160
ggggcaggca ccacctcccc taatgtg 2187

<210> 112
<211> 2020
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510232CB1

<400> 112
tgaaggctgg ttgttcagac tgagcttcct gcctgcctgt accccgccaa cagcttcaga 60
agaaggtgac tggtggctgc ctgaggaata ccagtgggca agagaattag catttctgga 120
gcatctgctg tctgagcagc ccctgggtgc gtccactttc tgggcacgtg aggttgggcc 180
ttggccgcct gagcccttga gttggtcact tgaaccttgg gaatattgag attatattct 240
cctgcctttt aaaaagatgg acttcagcag aaatctttat gatattgggg aacaactgga 300
cagtgaagat ctggcctccc tcaagttcct gagcctggac tacattccgc aaaggaagca 360
agaacccatc aaggatgcct tgatgttatt ccagagactc caggaaaaga gaatgttgga 420
ggaaagcaat ctgtccttcc tgaaggagct gctcttccga attaatagac tggatttgct 480
gattacctac ctaaacacta gaaaggagga gatggaaagg gaacttcaga caccaggcag 540
ggctcaaatt tctgcctaca ggttccactt ctgccgcatg agctgggctg aagcaaacag 600
ccagtgccag acacagtctg tacctttctg gcggagggtc gatcatctat taataagggt 660
catgctctat cagatttcag aagaagtgag cagatcagaa ttgaggtctt ttaagtttct 720
tttgcaagag gaaatctcca aatgcaaact ggatgatgac atgaacctgc tggatatttt 780
catagagatg gagaagaggg tcatcctggg agaaggaaag ttggacatcc tgaaaagagt 840
ctgtgcccaa atcaacaaga gcctgctgaa gataatcaac gactatgaag aattcagcaa 900
aggggaggag ttgtgtgggg taatgacaat ctcggactct ccaagagaac aggatagtga 960
atcacagact ttggacaaag tttaccaaat gaaaagcaaa cctcggggat actgtctgat 1020
catcaacaat cacaattttg caaaagcacg ggagaaagtg cccaaacttc acagcattag 1080
ggacaggaat ggaacacact tggatgcagg tacagtagaa cccaaaagag aaaagtaaaa 1140
tatttcttat gcctattttt ttttaaatca aaagggagag aacaaaagct ataccaaaag 1200
ggccatgttt caagaaaatg gatttaaaca tatttccctg tggaggggct ttgaccacga 1260
cctttgaaga gcttcatttt gagatcaagc cccacgatga ctgcacagta gagcaaatct 1320
atgagatttt gaaaatctac caactcatgg accacagtaa catggactgc ttcatctgct 1380
gtatcctctc ccatggagac aaaggcatca tctatggcac tgatggacag gaggcccccca 1440
tctatgagct gacatctcag ttcactggtt tgaagtgccc ttcccttgct ggaaaaccca 1500
aagtgttttt tattcaggct tgtcaggggg ataactacca gaaaggtata cctgttgaga 1560

ctgattcaga ggagcaaccc tatttagaaa tggatttatc atcacctcaa acgagatata 1620
tcccggatga ggctgacttt ctgctgggga tggccactgt gaataactgt gtttcctacc 1680
gaaaccctgc agagggaacc tggtacatcc agtcactttg ccagagcctg agagagcgat 1740
gtcctcgagg cgatgatatt ctcaccatcc tgactgaagt gaactatgaa gtaagcaaca 1800
aggatgacaa gaaaaacatg ggaaacaga tgcctcagcc tactttcaca ctaagaaaaa 1860
aacttgtctt cccttctgat tgatggtgct attttgtttg ttttgttttg ttttgttttt 1920
ttgagacaga atctcgctct gtcgcccagg ctggagtgca gtggcgtgat ctcggctcac 1980
cgcaagctcc gcctcccggg ttcaggccat tctcctgcct 2020

<210> 113
<211> 2167
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510233CB1

<400> 113
tgaaggctgg ttgttcagac tgagcttcct gcctgcctgt accccgccaa cagcttcaga 60
agaaggtgac tggtggctgc ctgaggaata ccagtgggca agagaattag catttctgga 120
gcatctgctg tctgagcagc cctgggtgc gtccactttc tgggcacgtg aggttgggcc 180
ttggccgcct gagcccttga gttggtcact tgaaccttgg gaatattgag attatattct 240
cctgcctttt aaaaagatgg acttcagcag aaatctttat gatattgggg aacaactgga 300
cagtgaagat ctggcctccc tcaagttcct gagcctggac tacattccgc aaaggaagca 360
agaacccatc aaggatgcct tgatgttatt ccagagactc caggaaaaga gaatgttgga 420
ggaaagcaat ctgtccttcc tgaaggagct gctcttccga attaatagac tggatttgct 480
gattacctac ctaaacacta gaaaggagga gatggaaagg gaacttcaga caccaggcag 540
ggctcaaatt tctgcctaca ggttccactt ctgccgcatg agctgggctg aagcaaacag 600
ccagtgccag acacagtctg tacctttctg gcggagggtc gatcatctat taataagggt 660
catgctctat cagatttcag aagaagtgag cagatcagaa ttgaggtctt ttaagtttct 720
tttgcaagag gaaatctcca aatgcaaact ggatgatgac atgaacctgc tggatatttt 780
catagagatg gagaagaggg tcatcctggg agaaggaaag ttggacatcc tgaaaagagt 840
ctgtgcccaa atcaacaaga gcctgctgaa gataatcaac gactatgaag aattcagcaa 900
agagagaagc agcagcctgg aaggaagtcc tgatgaattt tcaaatgggg aggagttgtg 960
tggggtaatg acaatctcgg actctccaag agaacaggat agtgaatcac agactttgga 1020
caaagtttac caaatgaaaa gcaaacctcg ggatactgt ctgatcatca acaatcacaa 1080
ttttgcaaaa gcacgggaga aagtgcccaa acttcacagc attagggaca ggaatggaac 1140
acacttggat gcaggggctt tgaccacgac ctttgaagag cttcattttg agatcaagcc 1200
ccacgatgac tgcacagtag agcaaatcta tgagattttg aaaatctacc aactcatgga 1260
ccacagtaac atggactgct tcatctgctg tatcctctcc catggagaca agggcatcat 1320
ctatggcact gatggacagg aggcccccat ctatgagctg acatctcagt tcactggttt 1380
gaagtgccct tcccttgctg gaaaacccaa agtgtttttt attcaggctt gtcaggggga 1440
taactaccag aaaggtatac ctgttgagac tgattcagag gagcaaccct atttagaaat 1500
ggatttatca tcacctcaaa cgagatatat cccggatgag gctgactttc tgctggggat 1560
ggccactgtg aataactgtg tttcctaccg aaaccctgca gagggaacct ggtacatcca 1620
gtcactttgc cagagcctga gagagcgatg tcctcgaggc gatgatattc tcaccatcct 1680
gactgaagtg aactatgaag taagcaacaa ggatgacaag aaaaacatgg ggaaacagat 1740
gcctcagcct actttcacac taagaaaaaa acttgtcttc ccttctgatt gatggtgcta 1800
ttttgtttgt tttgttttgt tttgtttttt tgagacagaa tctcgctctg tcgcccaggc 1860
tggagtgcag tggcatgatc tcggctcacc gcaagctccg cctcccgggt tcaggccatt 1920
ctcctgcctc agcctcccga gtagctggga ctacaggggc ccgccaccac acctggctaa 1980
ttttttaaaa atatttttag tagagacagg gtttcactgt gttagccagg gtggtcttga 2040
tctcctgacc tcgtgatcca cccacctcgg cctcccaaag tgctgggatt acaggcgtga 2100
gccaccgcgc ctggccgatg gtactattta gatataacac tatgtttatt tactaatttt 2160
ctagatt 2167

<210> 114
<211> 1706
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 7510304CB1

<400> 114
ctcgagccga ttcggctcga gcccggcact acctcgaggc tccgcccccа cctgctggac 60
cccagggtcc caccctggcc caggaggtca gccagggaat cattaacaag aggcagtgac 120
atggcgcaga aggagggtgg ccggactgtg ccatgctgct ccagacccaa ggtggcagct 180
ctcactgcgg ggaccctgct acttctgaca gccatcgggg cggcatcctg ggccattggt 240
gagagcggtt cctgtgcctc tgacctcccc tggcccctgc agcctccagc ctgcctgccc 300
tcacccctcc cttctctgca gtggctgttc tcctcaggag tgaccaggag ccgctgtacc 360
cagtgcaaga cgtggcggct gctgtgctcc tcgcgctcca acgccagggt agccggactc 420
agctgcgagg agatgggctt cctcagggca ctgacccact ccgagctgga cgtgcgaacg 480
gcgggcgcca atggcacgtc gggcttcttc tgtgtggacg aggggaggct gccccacacc 540
cagaggctgc tggaggtcat ctccgtgtgt gattgcccca gaggccgttt cttggccgcc 600
atctgccaag actgtggccg caggaagctg cccgtggacc gcatcgtggg aggccgggac 660
accagcttgg gccggtggcc gtggcaagtc agccttcgct atgatggagc acacctctgt 720
gggggatccc tgctctccgg ggactgggtg ctgacagccg cccactgctt cccggagcgg 780
aaccgggtcc tgtcccgatg gcgagtgttt gccggtgccg tggcccaggc ctctccccac 840
ggtctgcagc tgggggtgca ggctgtggtc taccacgggg gctatcttcc ctttcgggac 900
cccaacagcg aggagaacag caacgatatt gccctggtcc acctctccag tcccctgccc 960
ctcacagaat acatccagcc tgtgtgcctc ccagctgccg gccaggccct ggtggatggc 1020
aagatctgta ccgtgacggg ctggggcaac acgcagtact atggccaaca ggccggggta 1080
ctccaggagg ctcgagtccc cataatcagc aatgatgtct gcaatggcgc tgacttctat 1140
ggaaaccaga tcaagcccaa gatgttctgt gctggctacc ccgagggtgg cattgatgcc 1200
tgccagggcg acagcggtgg tccctttgtg tgtgaggaca gcatctctcg gacgccacgt 1260
tggcggctgt gtggcattgt gagttggggc actggctgtg ccctggccca gaagccaggc 1320
gtctacacca aagtcagtga cttccgggag tggatcttcc aggccataaa gactcactcc 1380
gaagccagcg gcatggtgac ccagctctga ccggtggctt ctcgctgcgc agcctccagg 1440
gcccgaggtg atcccggtgg tgggatccac gctgggccga ggatgggacg tttttcttct 1500
tgggcccggt ccacaggtcc aaggacaccc tccctccagg gtcctctctt ccacagtggc 1560
gggcccactc agccccgaga ccacccaacc tcaccctcct gaccccсatg taaatattgt 1620
tctgctgtct gggactcctg tctaggtgcc cctgatgatg ggatgctctt taaataataa 1680
agatggtttt gattaaaaaa aaaaaa 1706

<210> 115
<211> 2742
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510461CB1

<400> 115
ggggcgtggg ccggcaggaa gatggcgaac gtggggctgc agttccaggc gagcgcgggg 60
gactcggacc cacagagccg gcccctgctg ctgctcgggc agctgcacca cctgcaccgc 120
gtgccctgga gccacgtccg cgggaagctg cagccccggg tcaccgagga gctctggcag 180
gctgccctga gcacgctcaa ccccaacccc acggacagct gtcccctcta cctgaactac 240
gccaccgtgg ctgccctgcc ctgcagggtg agccggcaca acagcccctc ggccgcccac 300
ttcatcacgc ggctggtgcg gacctgcctg ccgcccggag cgcatcgctg cattgtgatg 360
gtctgcgagc agccggaggt ctttgcttcc gcctgtgccc tggcccgggc cttcccgctg 420
ttcacccacc gctcaggtgc ctctcggcgc ttggagaaga agacggtcac cgtggagttt 480
ttcctggtgg gacaagacaa cggccggtg gaggtgtcca cattgcagtg cttagcgaat 540
gccacagacg gcgtgcggct agcagcccgc atcgtggaca caccctgcaa tgagatgaac 600
accgacacct tcctcgagga gattaacaaa gttggaaagg agctggggat catcccaacc 660
atcatccggg atgaggaact gaagacgaga ggatttggag gaatctatgg ggttggcaaa 720
gccgccctgc atcccccagc cctggccgtc ctcagccaca ccccagatgg agccacgcag 780
accatcgcct gggtgggcaa aggcatcgtc tatgacactg gaggcctcag catcaaaggg 840
aagactacca tgccggggat gaagcgagac tgcgggggtg ctgcggccgt cctgggggcc 900
ttcagagccg caatcaagca gggtttcaaa gacaacctcc acgctgtgtt ctgcttggct 960
gagaactcgg tggggcccaa tgcgacaagg ccagatgaca tccacctgct gtactcaggg 1020
aagacggtgg aaatcaacaa cacggatgcc gagggcaggc tggtgctggc agatggcgtg 1080
tcctatgctt gcaaggacct gggggccgac atcatcctgg acatggccac cctgaccggg 1140

gctcaggtga gtgctccctg gatccacccc ttagctgtag tcccagggaa ccccacccca 1200
ctcttgacct caaggtgggg aaggcagcgc gtgggggctg ccgtcaggaa gttttcataa 1260
ctgggcacga ggaggcagga ggagctgccc gtcaagcttg gtgtgggcag taccgaggcc 1320
caggctggat gcagagccgg gaaccgcctc ctgtgtgctt catggccatg gtcctttctg 1380
cctgtgtttt ttcttttttt ctcaaccgtc tcttttctgg ctcccttatt tctctgtctg 1440
cctcccggtc cctcttttgc cttgggtgtt tctctcctgc cgtcccgtcc acacgcttcc 1500
cgggttcctg cccgcccagg gcattgccac agggaagtac cacgccgcgg tgctcaccaa 1560
cagcgctgag tgggaggccg cctgtgtgaa ggcgggcagg aagtgtgggg acctggtgca 1620
cccgctggtc tactgccccg agctgcactt cagcgagttc acctcagctg tggcggacat 1680
gaagaactca gtggcggacc gagacaacag ccccagctcc tgtgctggcc tcttcatcgc 1740
ctcacacatc ggcttcgact ggcccggagt ctgggtccac ctggacattg ctgcaccggt 1800
gcatgctggt gagcgagcca caggcttcgg tgtggccctc ctgctggcgc tcttcggccg 1860
tgcctctgag gaccctctgc tgaacctggt gtccccactg ggctgtgagg tggatgtcga 1920
ggagggggac ctggggaggg actccaagag acgcaggctt gtgtgagcct cctgcctcgg 1980
ccctgacaaa cggggatctt ttacctcact ttgcactgat taattttaag caattgaaag 2040
attgcccttc atatgggttt tggtttgtct ttctggtcgt cagcgtggtg gtggaaacag 2100
ctgaagtttt aggagacagc ttagggtttg gtgcgggcca cggggagggg accggaagc 2160
gctggggctt gtttctgttt gttacttaca ggactgagac atcttctgta aactgctacc 2220
cctggggcct tctgcacccc ggggtgaggc ctcctgcctg cctggtgccc tgtcccagcc 2280
ccaggtcctg tgcagggcac ctgcgtggct gacagccagg ctcttactcc agccggggct 2340
gccagcgcat ccagccagcc cagccctgtg aaagatggag ctgacttgct gcaggggacc 2400
tgatttatag ggcaagagaa gtcacactcc ggcctctcag aattcacttg aggttcaatt 2460
aaatacagtc acaccgcccc ctcaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2520
aaaagaaaac aaaacacagc agcacccaca acacgaaaaa gaaaaaaaga aaaagaccag 2580
cacaccagca gtaaacggga gaaccaaaaa aacgggggc gggacgcaaa gagacgcacg 2640
cgcacagcag caagaaaaac gaacaacacg gcgggcgaac aacacaacaa cgagcgaaaa 2700
gcaccacaca acgagtgggg aacagaccaa gcacgaatgg ga 2742

<210> 116
<211> 1362
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510392CB1

<400> 116
cggaactttc gtctcgtggg tgcgaaaggt aaccgaagcg gctcaggaag gcagctgtca 60
ctgagcccct ggaacagagc gagagtatcg taagtaacca ggctcagccg gtttctcagg 120
ccgctctagt caaataaacc ataaagatca gactcgggct tcttcacttc cttctctccg 180
tggtttcgcc attagcttcc ggttccgggg aggggccgag ttttcttcga agatttgggg 240
ctccgcgata cagttaggat ggctgtagta cctctgctgt tgttgggggg tttgtggagc 300
gctgtgggag cgtccagcct gggtgtcgtt acttgcggct ccgtggtgaa gctactcaat 360
acgcgccaca acgtccgact gcactcacac gacgtgcgct atgggtcagg tagtgggcag 420
cagtcagtga caggtgtaac ctctgtggat gacagcaaca gttactggag gatacggggg 480
aagagtgcca cagtgtgtga gaggggaacc cccatcaagt gtggccagcc catccggctg 540
acacatgtca acactggccg aaacctccat agtcaccact tcacttcacc tctttctgga 600
aaccagttgc tctgtaaggt gatattgtga tcattttata gaggagaaaa caaaggctca 660
aaggatttac agaagaaggg atcaagttga gattcaagga agtgagtgct tttggtgagg 720
aaggtgaagg tgattatctg gatgactgga cagtgctctg taatggaccc tactgggtga 780
gagatggtga ggtgcggttc aaacactctt ccactgaggt actgctgtct gtcacaggag 840
aacaatatgg tcgacctatc agtgggcaaa aagaggtgca tggcatggcc cagccaagtc 900
agaacaacta ctggaaagcc atggaaggca tcttcatgaa gcccagtgag ttgttgaagg 960
cagaagccca ccatgcagag ctgtgaatct agaggctctg agccactgtt aacgcacaat 1020
gttcacagac atctgttgct gcctcacctt gggatccctg ccacaagttc cttgggcagt 1080
ggccatgtca ccattgagat gaagatatac aacagaaaat agtggctgtg tttggaagct 1140
tcagccctgc acatttgaac tagtcactct cccagacttg cgtgggtcag ttctttctga 1200
gtagaggact tgctggtaaa ggggcagatg ctttttatta gtactgataa aacaaactga 1260
gggaaacatc cctcttagct gggaaacttt tactcttcag gagcttggca tcatggactg 1320
ttaatgtatg tgattttccc cctattttct ctctccaaaa aa 1362

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☑ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ______________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.